VOL. 107 NO. TE2. MARCH 1981

TRANSPORTATION ENGINEERING JOURNAL OF ASCE

PROCEEDINGS OF
THE AMERICAN SOCIETY
OF CIVIL ENGINEERS

AMERICAN SOCIETY OF CIVIL ENGINEERS
FOUNDED 1852

AEROSPACE
AIR TRANSPORT
HIGHWAY
PIPELINE
URBAN TRANSPORTATION

VOL. 107 NO. TE2. MARCH 1981

TRANSPORTATION ENGINEERING JOURNAL OF ASCE

PROCEEDINGS OF
THE AMERICAN SOCIETY
OF CIVIL ENGINEERS

AMERICAN SOCIETY OF CIVIL ENGINEERS FOUNDED 1852

AEROSPACE
AIR TRANSPORT
HIGHWAY
PIPELINE
URBAN TRANSPORTATION

ISSN 0569-7891

CONTENTS

This Journal is published bimonthly by the American Society of Civil Engineers. Publications office is at 345 East 47th Street, New York, N.Y. 10017. Address all ASCE correspondence to the Editorial and General Offices at 345 East 47th Street, New York, N.Y. 10017. Allow six weeks for change of address to become effective. Subscription price to members is $12.50. Nonmember subscriptions available; prices obtainable on request. Second-class postage paid at New York, N.Y. and at additional mailing offices. TE.

TECHNICAL NOTES

Proc. Paper 16076

DISCUSSION

Proc. Paper 16070

---

*Discussion period closed for this paper. Any other discussion received during this discussion period will be published in subsequent Journals.

16120 FARES FOR METERED RIDE SHARING VEHICLES

KEY WORDS: Calculations; **Cost analysis;** Cost sharing; Distance; **Fares;** Metering; Meters; **Pricing; Taxicabs;** Taxicab usage; Time factors; Urban transportation

ABSTRACT: An analytical relationship is presented which expresses more accurately the metered taxi fare as a function of both the distance traveled and the time elapsed through meter incrementing. The analytical framework is used to derive further relationships among parameters of shared ride service. The application of such an approach is also evaluated regarding automated fare calculation for exclusive ride taxi and shared ride transit in advance of a trip (when the distance and time for the trip can be reliably estimated). In the application, reliable distance and time estimates are obtained from a data base originally generated from geographic and travel data files developed by metropolitan organizations, and continuously enhanced through trip data collection.

REFERENCE: Ghahraman, David E., Au, Tung, and Baumann, Dwight M.B., "Fare Calculation for Metered Ride Sharing Vehicles," *Transportation Engineering Journal*, ASCE, Vol. 107, No. TE2, **Proc. Paper 16120**, March, 1981, pp. 105-117

16125 LAND SUBSIDENCE AND EARTH FISSURING

KEY WORDS: Aerial photography; Arizona; **Fissures; Geological surveys; Geologic investigations; Geologic structures;** Groundwater lowering; **Land subsidence;** Monitoring; Transportation problems

ABSTRACT: Land subsidence and earth fissures in central Arizona are geologic hazards being monitored by aerial photography. Ground-water level declines of 200ft - 400ft (61m - 122m) in the past 30 yr have caused subsidence of 2ft - 12ft (0.6m - 3.7m) in several areas totaling hundreds of square miles. Earth fissures—tensional cracks—appear near the margins of these areas. Subsidence is continuing in these areas and appearing in others. The number of fissures mapped is also increasing. Both subsidence and fissures, once confined to the rural agricultural areas, are now found in suburban and even urban locations. Both have affected highways, railroads, and pipelines.

REFERENCE: Scott, Lewis E., "Land Subsidence and Earth Fissuring in Central Arizona," *Transportation Engineering Journal*, ASCE, Vol. 107, No. TE2, **Proc. Paper 16125**, March, 1981, pp. 119-125

16132 DESIGN AND ACCEPTABLE WIND ENVIRONMENT

KEY WORDS: **Aerodynamics;** Buildings; **Climatology; Comfort;** Microclimatology; Pedestrians; Pedestrian safety; **Wind;** Windblast; Windbreaks; Wind chill; Wind forces; Wind pressure; Wind speed; **Wind velocity**

ABSTRACT: The comfort and safety of pedestrians has been neglected by designers because few criteria exist on acceptable wind velocities and it is difficult to predict the climatic characteristics around proposed buildings. Available information on wind effects on pedestrian comfort and safety are summarized. The mechanical effects of wind on comfort are now better understood than the thermal effects of climate. Limiting values of wind speed become the criteria to determine whether a space is comfortable or safe over time, and allow judgments to be made about project acceptability. Microclimatic-prediction techniques are explored, as are procedures for determining the probability of a proposed pedestrain area being uncomfortable or unsafe.

REFERENCE: Arens, Edward A., "Designing for an Acceptable Wind Environment," *Transportation Engineering Journal*, ASCE, Vol. 107, No. TE2, **Proc. Paper 16132**, March, 1981, pp. 127-141

16122 DOWNGRADE SPEED AND HEAVY VEHICLES

KEY WORDS: Buses; Highways; **Speed; Speed changes;** Speed patterns; Speed studies; Topographical factors; Truck handling; Trucks; **Vehicles; Vehicle weight**

ABSTRACT: The speed characteristics of heavy vehicles on downgrades are analysed. The analyses were made by field observations of truck speeds on six rural two-lane downgrades. Loaded trucks considerably reduced their speeds at the beginning of a downgrade. The amount of reduction, termed the approach-speed gradient, was related to the length and slope of the downgrade, the second variable contributing exponentially to the increase in gradient. Buses and empty trucks increased their speeds at the downgrades studied. Using a previously developed model, which is based on delay measures and the speed distribution of trucks and passenger cars, passenger-car equivalents for downgrades are calculated and evaluated.

REFERENCE: Polus, Abishai, Craus, Joseph, and Grinberg, Izhak, "Downgrade Speed Characteristics of Heavy Vehicles," *Transportation Engineering Journal,* ASCE, Vol. 107, No. TE2, **Proc. Paper 16122**, March, 1981, pp. 143-152

16126 TIME AND ROUTE ASSIGNMENT

KEY WORDS: **Commuting costs;** Commuting patterns; Models; **Peak hour traffic; Route preferences; Routes; Time factors;** Time spent driving; Time studies; Traffic research; **Transportation studies**

ABSTRACT: The Destination Target Time (DTT) is the time a commuter wishes to arrive at his destination. How the DTT affects a commuter's choice of departure time from home and his route for travel during the morning peak period is examined. A model for predicting the combined temporal distribution and route assignment of traffic during this period is presented. A commuter attaches some perceived costs to late and early arrivals at his destination and also to delays in the system. The model is developed by assuming that a commuter selects both his route for travel and departure times from home in order to minimize his total cost. An approximate algorithm is then developed from this model for practical application purposes, and illustrative examples are also presented.

REFERENCE: Alfa, Attahiru Sule, "Time-Dependent Route Assignment of Peak Traffic," *Transportation Engineering Journal,* ASCE, Vol. 107, No. TE2, **Proc. Paper 16126**, March, 1981, pp. 153-165

16130 DYNAMIC PAVEMENT DEFLECTIONS

KEY WORDS: **Deflection;** Elastic theory; Overlays (pavements); **Pavement deflections; Pavement design;** Pavement overlays; Pavement tests; Structural design

ABSTRACT: In 1977 a methodology was developed to evaluate pavement performance using dynamic (Road Rater) deflections. Since then, additional research has resulted in modifying the procedures. The procedures presently used to evaluate flexible pavement structures are explored, and background information is included on various procedures used by others. A sample set of data is presented and evaluated. A key to the adequate design of an overlay for any pavement structure is to be able to determine reasonable values for design parameters that represent the condition of the existing pavement. A discussion is included on how the analyses of dynamic pavement deflections can be used to design overlays and to manage pavements.

REFERENCE: Sharpe, Gary W., Southgate, Herbert F., and Deen, Robert C., "Dynamic Pavement Deflections," *Transportation Engineering Journal,* ASCE, Vol. 107, No. TE2, **Proc. Paper 16130**, March, 1981, pp. 167-181

16141 ASPHALT EMULSION TREATED BASES

KEY WORDS: **Asphalt emulsions; Asphalts;** Curing; Indirect tensile strength tests; Mixtures; **Pavement bases; Pavements;** Resilience; Temperature effects

ABSTRACT: A laboratory investigation was performed to characterize cold mixed asphalt emulsion treated bases. Indirect tension, resilient modulus, Hveen and water damage techniques were used in the mixture evaluation. The effects of aggregate type and gradation, added moisture content, asphalt emulsion content, and temperature were determined. The mixture was characterized at different curing conditions in the field. The effect of temperature on the tensile and resilient characteristics of the mixture is apparent. In addition, curing increases both the tensile strength and the resilient modulus of the mixture. Reasonable correlations are found between the indirect tensile test and the resilient modulus test results. Vacuum saturation has some effect on the tensile and the resilient properties of the mixture. Other asphalt emulsion mixture properties such as density, retained moisture content, and air voids are evaluated

REFERENCE: Mamlouk, Michael S., and Wood, Leonard E., "Characterization of Asphalt Emulsion Treated Bases," *Transportation Engineering Journal,* ASCE, Vol. 107, No. TE2, **Proc. Paper 16141**, March, 1981, pp. 183-196

16143 MIDTOWN MANHATTAN SURFACE TRANSIT STUDY

KEY WORDS: Automobiles; New York; Parking; Pedestrians; Transportation models; **Transportation planning;** Transportation studies; **Transportation system costs;** Urban buses; Urban traffic flow; **Urban transportation**

ABSTRACT: Low cost capital improvement techniques can be used to improve the efficiency of surface transportation systems as demonstrated in New York City. A project was undertaken to improve air quality in midtown Manhattan by redesigning vehicle usage and promoting transit. To assist in the evaluation of Midtown's surface transportation system, a conceptual model was developed with modal, operational and physical elements. All aspects of the movement of persons and vehicles were investigated and an overall circulation plan was developed. A number of projects were selected for demonstration to indicate the effectiveness of the various strategies in improving the efficiency of the transit system.

REFERENCE: Kraft, Walter H., and Schwartz, Samuel I., "Midtown Manhattan Circulation and Surface Transit Study and Demonstration Project," *Transportation Engineering Journal,* ASCE, Vol. 107, No. TE2, **Proc. Paper 16143**, March, 1981, pp. 197-212

16137 ALLOCATING PUBLIC TRANSIT COSTS

KEY WORDS: Allocation models; **Allocations; Cost allocation;** Cost analysis; **Public transportation;** Public transportation costs; Public transportation management; **Regional planning;** Rural transportation; **Transportation system costs**

ABSTRACT: A methodology was developed to aid in the selection of procedures to allocate regional transit costs among participants. The methodology provides a means for comparing a variety of allocation procedures, based on the equity of the cost allocation and the cost to use the procedure. Measures of equity were developed to allow comparison of each procedure with a prespecified "ideal" allocation. Data from a 100% sample of a regional transit service in Barnstable County, Mass., were used to illustrate the use of the procedure. Ten procedures were analyzed, each having different variables, data sampling frequencies, and weighting coefficients. The 10 procedures were compared to an allocation based on each participant's individual cost. Results showed that use of the methodology allowed elimination of several allocation procedures from further consideration.

REFERENCE: Male, James W., Collura, John, and Shuldiner, Paul W., "Allocating Public Transit Costs Among Participants," *Transportation Engineering Journal,* ASCE, Vol. 107, No. TE2, **Proc. Paper 16137**, March, 1981, pp. 213–225

16146 BURIAL DESIGN FOR FLOW CROSSINGS

KEY WORDS: **Buried cables; Buried pipes; Buried structures;** Cables; **Coastal engineering;** Estuaries; **Inlets (waterways);** Pipelines; Scouring; **Tidal waters**

ABSTRACT: Design criteria for burial depths (for submarine pipelines, cables, diffusors, etc.) across tidal rivers, estuaries and inlets are examined. For the case of shallow, alluvial tidal rivers and estuaries, design computation based upon real-time records is suggested. Numerical hydrodynamic models may be used to fine the instantaneous flow depths and velocities. The corresponding tractive stresses can then be computed for peak velocities by assuming a logarithmic velocity distribution. For tidal inlets, additional considerations must be given to the effects of waves on sediment transport near the seaward end of the inlet. A relationship for predicting the maximum scour depth in a new inlet as a function of the spring tidal prism is proposed. For inlets that are to be modified, the expected change in the maximum depth due to a change in the tidal prism can be computed.

REFERENCE: Graham, Donald Steven, and Mehta, Ashish J., "Burial Design Criteria for Tidal Flow Crossings," *Transportation Engineering Journal,* ASCE, Vol. 107, No. TE2, **Proc. Paper 16146**, March, 1981, pp. 227-242

TRANSPORTATION ENGINEERING JOURNAL

Fare Calculation for Metered Ride Sharing Vehicles

By David E. Ghahraman,[1] Tung Au,[2] F. ASCE, and Dwight M. B. Baumann[3]

(Reviewed by the Urban Transportation Division)

Introduction

The paratransit was first legitimized as a form of public transportation when in January 1975 it was briefly described in the guidelines for the apportionment program of Section 5 of the Urban Mass Transportation Act (2) and in September 1975, paratransit services were included in the Transportation Systems Management requirement of the joint FHWA-UMTA Planning Regulations (5). Paratransit is defined as any public transportation service that falls between private automobiles and line haul (fixed-route, fixed-schedule) transit service. The most significant characteristics of paratransit are door-to-door service, demand responsive scheduling, and shared use of vehicles.

Taxicab operators have traditionally played a dominant role in providing door-to-door demand responsive service. However, the problems of converting exclusive ride to shared ride service are complex, and fare structure is perhaps the most crucial factor influencing the success or failure of shared ride in a competitive market. A conventional electromechanical taxi meter records data for exclusive ride fares, but it cannot record adequate data for calculating shared ride fares. The zone fare concept has been used to alleviate customer fears of overcharging based upon circuitous routing as well as to provide a basis for shared ride fare calculation. With rare exceptions (Washington, D.C. being a notable one), however, zone fares exist primarily in small cities and rural

[1] Engr., Megatest Corp., Santa Clara, Calif.

[2] Prof. of Civ. Engrg. and Public Policy, Carnegie-Mellon Univ., Schenley Park, Pittsburgh, Pa. 15213.

[3] Prof. of Engrg. Design and Dir. of Center for Entrepreneurial Development, Carnegie-Mellon Univ., Schenley Park, Pittsburgh, Pa.; also President of Peoples Cab Co., Pittsburgh, Pa.

Note.—Discussion open until August 1, 1981. To extend the closing date one month, a written request must be filed with the Manager of Technical and Professional Publications, ASCE. Manuscript was submitted for review for possible publication on March 18, 1980. This paper is part of the Transportation Engineering Journal of ASCE, Proceedings of the American Society of Civil Engineers, ©ASCE, Vol. 107, No. TE2, March, 1981. ISSN 0569-7891/81/0002-0105/$01.00.

areas because the zone boundaries in large cities are not easily recognizable to the public. The zone fare concept works best when the users are on expense accounts or otherwise affluent, but proves to be least successful when the users are eager to verify the fares quoted by the drivers, but are incapable of doing so. Thus, the concept of incorporating travel distance and time in fare calculation, as represented by the conventional mechanical taxi meter, remains attractive to the users, providers and regulatory bodies as a fair and equitable solution whether the ride is exclusive or shared.

This paper presents an analytical relationship which expresses more accurately the metered taxi fare as a function of both the distance traveled and the time elapsed through meter incrementing. The analytical framework is used to derive further relationships among parameters of shared ride service. The paper also describes the application of such an approach to automated fare calculation for exclusive ride taxi and shared ride paratransit in advance of a trip when the distance and time for the trip can be reliably estimated. In the application described, reliable distance and time estimates are obtained from a data base originally generated from geographic and travel data files developed by metropolitan planning organizations, and continuously enhanced through trip data collection.

Analytical Framework for Metered Taxi Fare

Let F_0 be the fixed portion of the fare which is the initial charge for the flagdrop when a rider boards a taxicab, and let CN be the variable portion of the fare in which N = the number of times the meter trips itself during a ride; and C = the incremental rate or unit cost per meter tripping. The meter trips itself when either a specified distance increment δ or a specified time increment τ is reached first. The conventional metered taxi fare is therefore given by

$$F = F_0 + CN \qquad (1)$$

The vehicular traffic in congested urban areas is characterized by fast and slow movements as indicated by the trajectory of a vehicle with respect to position and time in Fig. 1. The instantaneous velocity v of a vehicle at any point is given by the slope of the trajectory dx/dt at that point. In the case of fast movement of the vehicle when $v > \delta/\tau$, the meter trips itself for every distance increment δ, and the point (x_n, t_n) representing, respectively, the position and time of the nth tripping of the meter is determined by the travel distance. In the case of slow movement when $v < \delta/\tau$, the meter trips itself for every time increment τ, and the point (x_n, t_n) is determined by the time lapsed. Thus, during a ride of total distance X and total time T, it can be expected that the meter reading is determined by a combination of distance segments and time segments. Then, the total fare F is equivalent to a linear combination of X and T, with suitable weighting factors A and B assigned to distance and time, respectively, i.e.:

$$F = F_0 + AX + BT \qquad (2)$$

in which A is expressed in cost per mile; and B in cost per hour. Then A

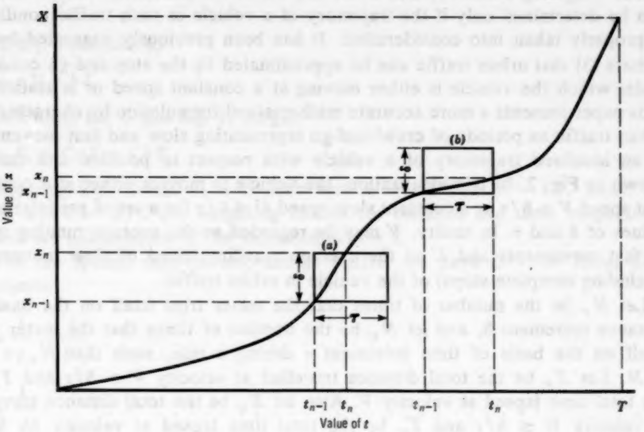


FIG. 1.—Trajectory of Vehicle Location with Respect to Distance and Time: (*a*) Fast Movement; (*b*) Slow Movement

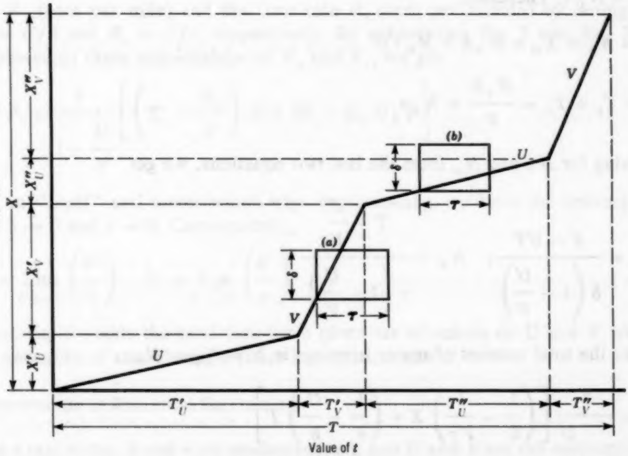


FIG. 2.—Idealization of Trajectory in Crawl-and-Go Traffic; (*a*) At Average Running Speed *V*; (*b*) At Average Crawling Speed *U*

and B may may be regarded as fare (or revenue) per vehicle mile and vehicle hour, respectively.

The weighting factors A and B are dependent on the traffic conditions, and can be determined only if the trajectory of a vehicle in such traffic conditions is properly taken into consideration. It has been previously suggested by the writers (3) that urban traffic can be approximated by the stop-and-go condition under which the vehicle is either moving at a constant speed or is stationary. This paper presents a more accurate mathematical formulation by characterizing urban traffic as periods of crawl-and-go representing slow and fast movements in an idealized trajectory of a vehicle with respect to position and time as shown in Fig. 2. In this idealization, the vehicle is moving either at a constant fast speed $V > \delta/\tau$ or a constant slow speed $U \leq \delta/\tau$ for a set of predetermined values of δ and τ. In reality, V may be regarded as the average running speed of fast movements and U as the average crawling speed of slow movements (including complete stops) of the vehicle in urban traffic.

Let N_V be the number of times that the meter trips itself on the basis of distance increment δ, and let N_U be the number of times that the meter trips itself on the basis of time increment τ during a ride, such that $N_V + N_U = N$. Let X_V be the total distance travelled at velocity $V > \delta/\tau$ and T_V be the total time lapsed at velocity V. Also let X_U be the total distance travelled at velocity $U \leq \delta/\tau$ and T_U be the total time lapsed at velocity U. Since X_V and T_V approximate the values of total distance and total time respectively for the largest integer N_V, and X_U and T_U approximate the corresponding values for the largest integer N_U, we note $X_V = N_V\delta$, $T_V = X_V/V$ and $X_U = T_U U$, $T_U = N_U\tau$. Therefore:

$$X = X_V + X_U = N_V\delta + N_U\tau U \quad \text{(3a)}$$

$$T = T_V + T_U = \frac{N_V\delta}{V} + N_U\tau \quad \text{(3b)}$$

Solving for N_V and N_U from the last two equations, we get

$$N_V = \frac{X - UT}{\delta\left(1 - \dfrac{U}{V}\right)}; \quad N_U = \frac{T - \dfrac{X}{V}}{\tau\left(1 - \dfrac{U}{V}\right)} \quad \text{(3c)}$$

Since the total number of meter trippings is $N = N_U + N_V$:

$$N = \frac{1}{1 - \dfrac{U}{V}}\left[\left(\frac{1}{\delta} - \frac{1}{V\tau}\right)X + \left(\frac{1}{\tau} - \frac{U}{\delta}\right)T\right] \quad \text{(4)}$$

Let $a = \dfrac{1}{1 - \dfrac{U}{V}}\left(\dfrac{1}{\delta} - \dfrac{1}{V\tau}\right)$ . . . . . (5a)

$$b = \frac{1}{1 - \dfrac{U}{V}}\left(\frac{1}{\tau} - \frac{U}{\delta}\right) \qquad (5b)$$

then Eq. 4 reduces to

$$N = aX + bT \qquad (6)$$

Substituting Eq. 6 into Eq. 1, we get

$$F = F_0 + CaX + CbT \qquad (7)$$

Comparing Eqs. 2 and 7, it is seen that $A = Ca$ and $B = Cb$. Thus, $a = A/C$ and $b = B/C$ are weighting factors for distance and time, respectively, which are independent of the incremental rate C.

For the special case of stop-and-go traffic, we have $U = 0$. Then, Eq. 5 becomes

$$a = \frac{1}{\delta} - \frac{1}{V\tau} \qquad (8a)$$

$$b = \frac{1}{\tau} \qquad (8b)$$

which are identical to the results previously obtained (5). Thus, the total fare can still be calculated by Eq. 7.

It is sometimes convenient to express the total fare in terms of the distance rate R_x (cost per mile) and the time rate R_t (cost per minute) by noting that $R_x = C/\delta$ and $R_t = C/\tau$, respectively. By substituting Eq. 5 into Eq. 7 and by observing these relationships of R_x and R_t, we get

$$F = F_0 + \frac{1}{1 - \dfrac{U}{V}}\left[\left(R_x - \frac{R_t}{V}\right) X + (R_t - R_x U)\, T\right] \qquad (9)$$

For an "ideal" taxi meter which trips continuously, we have the limiting case that $\delta \rightarrow 0$ and $\tau \rightarrow 0$. Consequently,

$$R_x = \lim_{C,\delta \rightarrow 0}\left(\frac{C}{\delta}\right); \quad R_t = \lim_{C,\tau \rightarrow 0}\left(\frac{C}{\tau}\right) \qquad (10)$$

Then, Eq. 9 yields the total fare for a given set of values of U and V without the necessity of considering a finite incremental rate C.

Interpretation of Analytical Relationship

In a taxi meter, δ and τ are predetermined, and U and V are the only unknown quantities associated with the weighting factors a and b in Eq. 5. It is therefore appropriate to examine U and V in detail. In the idealization of the trajectory of vehicle location, the vehicle has been assumed to move either at a constant fast speed $V > \delta/\tau$ or a constant slow speed $U \leq \delta/\tau$. In reality, there may be many fast movements of varying speeds interposed by slow movements

of varying speeds (including complete stops). However, all fast movements or slow movements in a trip can be averaged to obtain the average running speed V and the average crawling speed U, respectively. In Fig. 2, e.g., a stretch of slow movements is followed by a stretch of faster movements, which is in turn followed by another round of slow and fast movements during the trip; thus, two stretches of U and V each are shown in the trajectory. Then, the total distance X_U traveled at the average crawling speed U is the sum of X'_U and X''_U, and the total time elapsed T_U while traveling at the average crawling speed U is the sum of T'_U and T''_U. Similarly, the total distance X_V traveled at the average speed V is the sum of X'_V and X''_V and the total time elapsed T_V while traveling at the average running speed V is the sum of T'_V and T''_U.

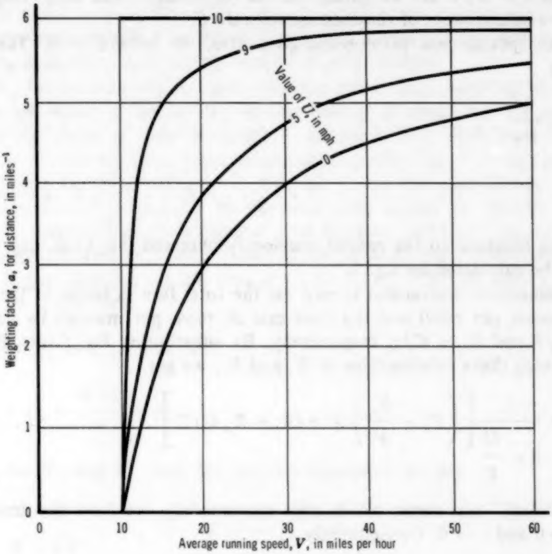


FIG. 3.—Variation of Weighting Factor for Distance with Vehicle Speeds

If the travel speed data are collected for a large number of trips in urban traffic, the average speeds U and V computed on the basis of such data will provide reasonable approximations for fare calculation.

To evaluate the effects of U and V on the weighting factors a and b for the predetermined values of δ and τ, let $\delta = 1/6$ mile (0.27 km) and $\tau = 1/60$ h. Then, the variations of weight factors for distance and time with vehicle speeds are shown in Figs. 3 and 4, respectively. Note that $\delta/\tau = 10$ mph (16.1 km/h) has been used as the dividing point for U and V, i.e., $V > 10$ mph and $U \leq 10$ mph. Then, for $U = 5$ mph (8.05 km/h) and $V = 30$ mph (48.3 km/h) which are reasonably realistic values for urban traffic, it is found from Figs. 3 and 4 that $a = 4.8$ and $b = 36$.

Metered Fare for Shared Ride

The conventional taximeter, when adapted to shared ride paratransit, would accumulate passenger fares on the basis of discounted distance and time rates. The attraction of shared ride is fare reduction, and its profitability is increased revenue per vehicle mile and vehicle hour. The objectives of reducing fare while increasing revenue when converting from exclusive ride taxi service to shared ride paratransit are affected by fare rate discount as well as by vehicle occupancy and route deviation. Simplified relationships representing these effects can be derived by considering the average values of vehicle occupancy and route deviation and by assuming the equivalence of distance averaging and

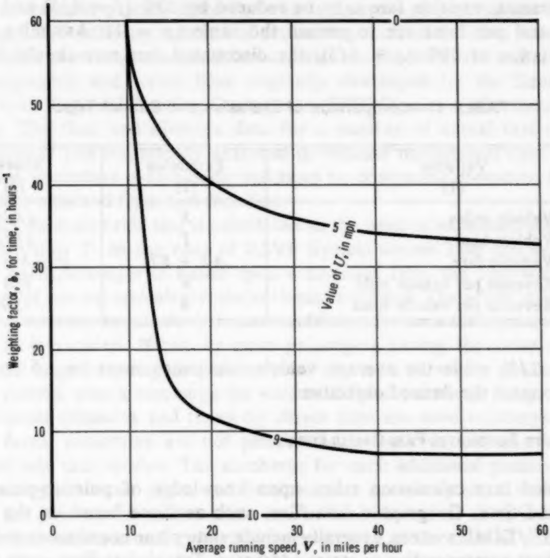


FIG. 4.—Variation of Weighting Factor for Time with Vehicle Speeds

time averaging operations. (This assumption is satisfied in the long run when the ratio of distance to time approaches a constant velocity characteristic of the service area.)

Let s denote average vehicle occupance. Then, s is the ratio of passenger distance (or time) to vehicle distance (or time). For exclusive rides, vehicle and passenger distances (or times) are equivalent. Let q be the average route deviation multiplier. The distance X traveled for an exclusive ride is equivalent to passenger distance qX and vehicle distance qX/s in the case of shared ride for the same trip. Similarly, time T for exclusive service is equivalent to passenger time qT and vehicle time qT/s for shared service. These definitions and Eq.

2 are used to develop Table 1 in which r represents the fraction of fare rate or discounted fare rate.

From Table 1, shared ride variable fare is expressed as discounted exclusive ride variable fare by

$$f = qr \qquad (11)$$

in which $0 < r < 1$; and $q > 1$. The ratio of shared ride revenue per vehicle distance (or time) to the corresponding quantity for exclusive ride is

$$p = rs \qquad (12)$$

in which $s > 1$.

To illustrate the use of Eqs. 11 and 12, suppose that in conversion to shared ride paratransit, variable fare is to be reduced by 20% ($f = 4/5$) and revenue per mile and per hour are to remain the same ($p = 1$). Assuming average route deviation of 20% ($q = 6/5$), the discounted fare rate should be set at

TABLE 1.—Comparison of Exclusive and Shared Trips

Variable (1)	Exclusive (2)	Shared (3)
Vehicle miles	X	qX/s
Vehicle hours	T	qT/s
Variable fare	$AX + BT$	$qr(AX + BT)$
Revenue per vehicle mile	A	rsA
Revenue per vehicle hour	B	rsB

67% ($r = 2/3$) while the average vehicle occupancy must be 1.5 ($s = 3/2$) in order to meet the desired objectives.

Application to Automated Fare Calculation

Automated fare calculation relies upon knowledge of point-to-point travel distance and time. Geographic data files, such as those based on the Census Bureau GBF/DIME system, generally include state plane coordinates and traffic zone numbers corresponding to street addresses. Travel data files, such as those developed by metropolitan planning organizations provide distance and time between zone centroids as well as centroid coordinates. Thus, for specified trip origin and destination, trip distance X and time T can be estimated and used as in Eq. 7 to calculate trip fare.

Geographic and travel data files do not in general provide time-sensitive data. Such files, however, can be used to generate the original information in an adaptive data base. Through trip data collection, the data base as well as fare calculation parameters can be continuously updated and eventually stratified with respect to time.

Trip data collection can be done by means of taximeter counters which record travel distance X, travel time T, number of meter trippings due to distance N_V, number of meter trippings due to N_U and the number of passengers onboard. Collected values of X and T are used to adapt the data base to travel conditions.

Sets of values of X, T, N_U, and N_V can be used to solve for U and V in Eqs. 3 and to adapt fare calculation to travel conditions as well. In addition, data collected from actual trips can keep running averages of vehicle occupancy and route deviation. Such averages find use through Eqs. 11 and 12 in metering shared ride trips in the event of fare calculation computer breakdown.

An experimental computer-based system to compute automatically and display, in advance, fares and estimated trip times for exclusive and share taxi services has been developed at Carnegie-Mellon University (1). The complete hardware and software packages for providing such capabilities are collectively termed the Ride Shared Vehicle Paratransit (RSVP) system. The system includes central and vehicle hardware which supports two-way digital communication between the control center minicomputer and the taximeters. The communication link is used to transmit precalculated fares to taximeters as well as to collect trip data. The RSVP software consists of processing routines for data base management, vehicle communication, and fare calculation. The data base was constructed from geographic and travel files originally developed by the Southwestern Pennsylvania Regional Planning Commission (SPRPC) as an aid for transportation planning. The time and distance data for a number of actual taxi trips have been collected and statistically analyzed to validate the original data. The fare calculation procedure is also being validated by comparing computed fares with meter fares obtained from taxi manifest.

Fare for exclusive ride taxi is calculated on the basis of estimated trip distance X and trip time T. In the case of RSVP System shared ride taxi service, the fare for each passenger is based upon a discount from the fare which would be charged if normal taxicab service has been requested. Therefore, overcharging because of excessive route deviation or insufficient fare rate discount as expressed by Eq. 11 is avoided. If two or more passengers having the same origin and destination request group taxicab service, the total fare is a discounted fare for one person, plus a surcharge for each additional person in the group. Since point-to-point distances and times for direct trips are used to compute normal taxicab fares, customers are not penalized for route deviations which occur in shared ride taxi service. The surcharge for each additional passenger in any group having the same origin and destination is necessary since in shared taxi service the number of vacant seats is a critical parameter (i.e., it determines whether the vehicle is still potentially shareable). At the time of a request for service, the customer is informed of the least expensive class of service, the fare, the estimated time before a vehicle would be available, and the estimated travel time for exclusive taxicab service.

Regulatory Issues in Implementation

The RSVP system has been implemented for preliminary testing within an institutional and regulatory environment, utilizing the Peoples Cab Company of Pittsburgh owned by the Center for Entrepreneurial Development, a nonprofit corporation affiliated with Carnegie-Mellon University. As a common carrier of passengers by motor vehicles, the company is regulated by the Pennsylvania Public Utility Commission (PUC). Prior to performing any experiments with the RSVP System in an actual operating environment, the company was required to obtain regulatory approval both for the general concept of shared ride service

and for the fare calculation and display mechanisms of the RSVP System.

In Pennsylvania, taxi group riding was prohibited except in specific circumstances until the recent energy crisis. In 1974, the Pennsylvania PUC issued an order that somewhat eased restrictions on group riding, although the associated fare calculation procedure was cumbersome and inequitable. With this PUC order as a precedent, Peoples Cab Company proposed in 1976 a tariff based upon the RSVP System, which modified the commission's fare calculation procedure and further relaxed the restrictions on group riding.

In contrast with most other taxi companies in Pennsylvania, Peoples Cab Company chose to file a shared ride tariff under its existing Certificate of Public Convenience, rather than to apply for a new "Special Operations" Certificate for paratransit services. This approach has particular significance in Allegheny County, Pa., since Port Authority Transit (PAT, the local mass transit agency) has both the authority and the responsibility to provide an integrated transportation system for the area. Although taxi service is specifically excluded by its enabling legislation, PAT has taken the position that it has the authority to regulate shared ride services if such services are provided in vehicles other than sedans. Furthermore, a number of social service agencies in the County are operating vehicles under the auspices of Section 16 (b) 2 of the Urban Mass Transportation Act. Many of these agencies compete directly with existing private taxi companies and have allegedly caused substantial reductions in those companies' revenues. Since the lower rates offered by social service agencies for their clients are only possible because of subsidies in the form of vehicles, these practices raise the issue of efficiency versus equity. Resolution of these jurisdictional disputes related to shared ride services will clearly have important ramifications for urban and suburban transportation, and the fact that a duly certificated taxi company has an approved shared ride tariff (rather than a temporary certificate for a new service) should help insure that taxi operators are given a fair opportunity to participate in the growing paratransit market.

Specific provisions of the Peoples Cab Company tariff define group taxicab service as nonexclusive, door-to-door call or demand service in which persons having different origins and destinations share a vehicle. Since such ride sharing inevitably involves route deviations, each individual's fare is first computed as if the trip would be an exclusive taxi trip. Thus, in terms of time and distance used to compute fares, there is no penalty imposed on a customer due to route deviations. After the basic point-to-point fare is computed, it is discounted depending upon the degree of advance notification given to the company by the customer. In addition, the tariff contains a unique dedicated vehicle provision which permits the company to negotiate with organizations individually to determine the charge for services, subject to a maximum rate per mile or per hour. Through the regulatory approval of group taxicab service proposed by Peoples Cab Company, it is now possible to offer shared ride, subscription service and service to affinity group by a certificated common carrier of passengers in Pennsylvania.

The change in taxi-paratransit regulations is a slow process because the regulatory agency is embedded in a political and economic environment and its decision making process encompasses many possible trade-offs involving numerous parameters and interest groups. Consequently, the experiments with

the RSVP System in an actual operating environment may be viewed as the change agents. The regulatory changes proposed by Peoples Cab Company as a result of the experiments in using the RSVP System are incorporated in the latest (September 1980) version of the regulations for motor carriers of passengers issued by Pennsylvania PUC (4). Since Pennsylvania is one of the most populous states having statewide transportation regulation, the regulatory issues resolved in Pennsylvania should serve as useful guidelines for other areas throughout the United States.

Operating Problems

A number of operational problems have inhibited the growth of shared taxi systems. Such problems arise primarily from the lack of adequate controls over cash transactions that occur in vehicles. From the customer's point of view, a fare calculation and collection system should not allow a driver to influence the fare for any reason. From management's perspective, a fare calculation and collection system should permit management to know of every trip and to charge appropriate fares according to the tariff granted by regulatory authorities. Furthermore, both management and regulatory authorities require access to accurate and complete operational data for use in determining fare structures and assessing appropriate taxes and fees.

The introduction of the RSVP System and the ensuring regulatory changes oy the Pennsylvania PUC have given a significant boost toward the removal of some of these operating problems. Specifically, the new PUC regulations (4) include the following statements (italics added):

1. Under the scheme of classification of common carriers, call-or-demand service is now defined as "local common carrier service for passengers, rendered on *either an exclusive or a nonexclusive basis*, where the service is characterized by the fact that passengers normally hire the vehicle and its driver either by telephone or by hail, or both." [Section 29.13 (2)]

2. With regard to taxi meters, it is now stated that "no meter shall be operated from any drive other than the transmission of such vehicles *unless some other method is, upon petition, specifically approved by the commission*." [Section 29.314 (b) (3)]

3. For a call-or-demand service under tariffs authorizing both exclusive and nonexclusive service, "the dispatcher shall, if requested by the customer, quote the customer *the estimated fare for the customer's trip as priced under both of these two alternative services*, considering the number of people in the customer's travelling group; and the dispatcher shall explain to the customer, if necessary, the difference in these two types of service." [Section 29.316 (d)]

These new regulations will allow the use of a fairer and more equitable fare structure in providing paratransit services by the call-or-demand service certificated carriers. It remains to be determined what constitutes a fair and equitable fare structure with regard to the difference in fares for exclusive and nonexclusive rides. In conducting experimental operations of the RSVP System in the real world environment, the discount rates for nonexclusive and group services have

been deliberately kept small so that the cab company will not be inundated with requests for nonexclusive service. The fares set by the market will be greatly influenced by the public policies in subsidizing public transportation. If user subsidies are given to the elderly, handicapped, and economically disadvantaged in lieu of the subsidies to the providers of paratransit service, it is more likely that the service will be more efficient as well as more equitable.

The acceptance of the RSVP System computer-based fare calculation procedures by the regulatory agency will permit a wide variety of marketing and fare structure experiments, utilizing monitoring and analysis capabilities that have previously not been available. These same monitoring and analysis capabilities should also aid taxi companies to significantly improve operating efficiencies, and should provide the paratransit industry with overall improvements in information processing capabilities. Eventually, travel time and travel distance data automatically recorded by the RSVP System for all trips undertaken will provide valuable information for updating the computer-based data file for fare calculation. The availability of appropriate technologies which are compatible with regulatory requirements and are priced within the reach of small operators will encourage maximum feasible participation of private enterprise in fulfilling local urban transportation needs.

Conclusion

Although the application of the automated fare calculation described in this paper has been limited so far to an operating taxicab company in Pittsburgh, the analytical framework is applicable to all metered ride sharing vehicles. Thus, the improvement of an analytical framework for metered fare calculation is useful in providing a fairer and more equitable fare structure for shared ride paratransit services.

Acknowledgment

The ideas in this paper were developed in connection with a project supported by a UMTA Service and Demonstration Grant No. PA-06-0048, United States Department of Transportation.

Appendix I.—References

1. Au, T., and Baumann, D. M. B., "Ride Shared Vehicle Paratransit System," *Report No. DOT-TST-76-84,* National Technical Information Service, Springfield, Va., 1977.
2. "Capital and Operating Assistance Formula Grants, Interim Guidelines and Procedures," *Federal Register,* Urban Mass Transportation Administration, U.S. Department of Transportation, Vol. 40, No. 8, Jan., 1975, pp. 2534–2557.
3. Ghahraman, D., Au, T., and Baumann, D. M. B., "Analysis of Metered Taxi Fares," *Journal of Transportation Engineering,* ASCE, Vol. 101, No. TE4, Proc. Paper 11735, Nov., 1975, pp. 807–816.
4. "Motor Carriers of Passengers," *Pennsylvania Bulletin,* 52 PA. Code Chapters 23 and 29, Pennsylvania Public Utility Commission, Vol. 10, No. 36, Sept., 1980, pp. 3614–3624.
5. "Transportation Improvement Program," *Federal Register,* Federal Highway Administration and Urban Mass Transportation Administration, U.S. Department of Transportation, Vol. 40, No. 181, Sept., 17, 1975, pp. 42976–42984.

APPENDIX II.—NOTATION

The following symbols are used in this paper:

A = weight for distance, in cost per mile;
a = weight for distance independent of incremental cost, per mile;
B = weight for time, in cost per minute;
b = weight for time independent of increment cost, per minute;
C = incremental cost;
F = total fare;
F_c = variable portion of fare;
F_0 = fixed portion of fare;
f = fraction of variable portion of fare;
N = total number of meter trippings;
N_U = number of meter trippings based on time increment τ;
N_V = number of meter trippings based on distance increment δ;
p = ratio of revenues per distance or time;
q = average route deviation multiplier;
R_t = time rate of fare, in cost per minute;
R_x = distance rate of fare, in cost per mile;
r = fraction of fare rate;
s = average vehicle occupancy;
T = total time elapsed;
T_U = total time elapsed at velocity $U \leq \delta/\tau$;
T_V = total time elapsed at velocity $V > \delta/\tau$;
t_n = time at nth meter tripping;
U = average running speed of slow movements ($\leq \delta/\tau$);
V = average running speed of fast movements ($> \delta/\tau$);
v = instantaneous velocity;
X = total distance traveled;
X_U = total distance traveled at velocity $U \leq \delta/\tau$;
X_V = total distance traveled at velocity $V > \delta/\tau$;
x_n = position of vehicle at nth meter tripping;
δ = distance increment preset in meter; and
τ = time increment preset in meter.

TRANSPORTATION
ENGINEERING JOURNAL

Land Subsidence and Earth Fissuring in Central Arizona[a]

By Lewis E. Scott[1]

(Reviewed by the Aerospace Division)

Introduction

Land subsidence and related earth fissuring have been known in Arizona for 50 yr. They have occurred in several of the deep alluvial basins in the central part of the state (see Fig. 1). Their relationship to large-scale withdrawal of ground water was suspected in the early 1960s, and generally confirmed by studies made in the last 10 yr. Land subsidence, or simply subsidence, as described in this paper, is the general and gradual sinking of a large area covering part or most of a large valley or basin area. Typically, the outline of the sinking area reflects the shape of the basin and is roughly circular, oval-shaped, or elongated. It is dish-shaped in the cross section, with a few tenths of a foot of settlement at the edges increasing to a maximum of 10 ft–12 ft (3 m–3.7 m) of settlement near the center. Earth fissures are tensional cracks occurring at various locations within the subsiding area; usually near its margin or near the mountain front forming the edge of the basin. It is only recently that the seemingly random, unpredictable locations of the fissures could be explained with some confidence.

In Arizona, earth fissures are nearly vertical cracks 0.5 in.–0.75 in. (19 mm) in width (see Fig. 2). They are rarely seen on the ground surface as narrow cracks, this width being due to erosion. Irrigation water or sheet flow drainage from rain storms quickly widens the narrow crack by erosion and slumping to a gulley 2 ft–20 ft (0.6 m–6.1 m) in width, and half as deep. It is these large features that are seen in the field, and are known as fissure gulleys to differentiate them from the actual fissure tensional crack (see Fig. 3).

[a]Presented at the April 14–18, 1980, ASCE Annual Convention and Exposition, held at Portland, Oreg.

[1]Sr. Geological Engr., Highways Div., Arizona Dept. of Transportation, Material Services, 1745 West Madison Street, Phoenix, Ariz. 85007.

Note.—Discussion open until August 1, 1981. To extend the closing date one month, a written request must be filed with the Manager of Technical and Professional Publications, ASCE. Manuscript was submitted for review for possible publication on April 22, 1980. This paper is part of the Transportation Engineering Journal of ASCE, Proceedings of the American Society of Civil Engineers, ©ASCE, Vol. 107, No. TE2, March, 1981. ISSN 0569-7891/81/0002-0119/$01.00.

A fissure usually consists of a single crack, but may open into a branched or even braided pattern up to a few inches in total width. The crack is usually filled with loose soil, often of a lighter color, which has migrated downward due to water action. These cracks have been traced visually in test pits to a depth of 16 ft (5 m) and probably extend to a depth of several hundred feet (100 m–200 m). They are thought to originate at depth, and migrate upward toward the ground surface. They are not geologic faults and do not exhibit

FIG. 1.—Location of Water-Level Declines and Earth Fissures in Central Arizona

lateral movement. Only a few of the several hundred observed have had any vertical displacement.

Geologic Origins

Central Arizona consists of a series of parallel mountain ranges separated by broad, flat, alluvium-filled basins. The mountains are of fault block origin forming the Basin-Range Geologic Provence which extends from Arizona northwestward through Nevada and eastern Oregon into eastern Washington. Each large basin in central Arizona contains a valley fill of clay, silt, sand, and gravel from 1,000 ft–2,000 ft (300 m–600 m) in depth. This material is generally permeable and forms a vast reservoir for ground water. Prior to about 1920, the average rainfall of approx 8 in. (200 cm) per year was able to replace what ground water was used and maintain a state of equilibrium. Since that time, pumpage has increased at a rapid rate creating an inbalance or "overdraft." The United States Geological Survey (USGS) records show the ground-water

surface has declined 200 ft–400 ft (60 m–120 m) in most of these basins since the later 1940s. Compaction-recorder data indicate long-term ground-water declines correspond to land subsidence. One site measured near Eloy between April 1965 and April 1974 recorded 1.34 ft (0.4 m) settlement in the upper 830 ft (253 m) of alluvium, 63% of the 2.13 ft (0.65 m) of subsidence measured at this site (6). The consolidate that occurs within the dewatered sediments is irreversible and reduces its capacity to store ground water by as much as 50%. Within the last decade both land subsidence and earth fissures have become recognized as geologic hazards in Arizona. Neither are unique to Arizona. Papers presented at the Third International Conference on Land Subsidence, held in

FIG. 2.—Recently-Formed Earth Fissure; 100-mm Cigarette in Center for Scale

Florida in 1978 (4), described land subsidence due to withdrawal of oil, gas, water, or collapse of underground openings in several areas of California, Texas, and Illinois in this country, in Holland, Norway, and Italy in Europe, and in Thailand and Taiwan in Asia. Fissures are known in Texas and California.

Aerial Mapping

While earth fissures are usually recognizable from the ground or from the air, land subsidence is not. Subsidence is gradual and may affect an area of

several hundred square miles. Its presence was first suspected and then confirmed in several areas in central Arizona as variations in the National Oceanic and Atmospheric Administration (NOAA) first-order level lines were plotted by USGS geologists and separately by Arizona Department of Transportation (ADOT) engineers. Requests for information and aerial photography between the two organizations indicated a common interest which developed into an Inter-Agency Study Group in the late 1960s. This group consisted of representatives of the USGS, ADOT, NOAA, United States Air Force, United States Bureau of Reclamation (USBR), and interested university and county officials. An inventory of known subsidence areas and earth fissures was begun using existing aerial

FIG. 3.—Fissure Gulley: Original Narrow Crack Opened by Water Flow and Slumping of Sides

photography supplied by the ADOT and the Air Force. The Highways Division of the ADOT has a fully-equipped photogrammetry and aerial mapping department with its own aircraft and cameras. Routine and periodic flights have been made along the principal highway routes since 1959 by this department at scales of 1 in. equals 1,000 ft (1:1 000), and 1 in. equals 2,000 ft (1:2 000). Individual prints and stereo pairs from these flights were used to plot the location of level-line bench marks and to monitor earth fissures. A series of photos taken at 3 yr-4 yr intervals shows the enlargement of existing fissures and the location or development of new fissures.

ADOT engineers and geologists were interested in this phenomenon at this time because of the large fissure system crossing Interstate Highway 10 near Picacho Peak which was causing a continuing maintenance problem (see Fig. 4). This fissure is unusual in that it exhibits vertical displacement. The area on both sides of the fissure is subsiding, but the north side is 2.5 ft–3.0 ft (0.76 m–0.9 m) lower than the south side. Holzer (1) considers this feature to be a fault related to a bedrock offset he interprets to be a fault. All other investigators (3,4,5,6) feel that its origin is ground-water withdrawal and that

FIG. 4.—Picacho Fissure System Crossing Interstate 10 in Foreground; Fissure is Over 9 mile (14.5 km) in length

FIG. 5.—Picacho Fissure Crossing Interstate 10 Showing Ramped Repaired Section; Note Vertical Displacement of Light Colored Curb

it is a fissure caused by the consolidation within the deep valley fill. The vertical displacement is corrected by long ramps on both eastbound and westbound roadways (see Fig. 5). On several occasions backhoe test pits or auger borings have been made on the fissure to check the possibility of continued settlement causing a void to form beneath the pavement. To date no problems of this type have developed. The Southern Pacific Railroad also has a constant maintenance problem at this and one other site due to fissures crossing their roadbed.

Fissure gulleying has exposed high pressure gas lines in the area, but has caused no damage.

Cracking of highway pavements by new fissures or continued development of older fissures has become frequent enough to be treated as a routine problem. Smaller cracks are filled and the larger cracks dug out and patched. The reaction to the presence of fissures in a proposed subdivision area or existing commercial area, however, is completely different; the matter is handled in total secrecy and prompt action is taken to remove all traces.

By the early 1970s special flights were made to map the fissures. The mapping was done in two stages. First, color infrared aerial photographs at scales of 1:24 000 and larger were made of all known fissure areas. A total of 144 aerial photographs were obtained by aircraft and camera provided by the ADOT and the film by the USGS. All identified fissures and probable fissure locations were transferred from this photography to 1:24 000-scale orthophotoquads furnished by the Arizona Resources Information System. Second, these fissures were checked and the entire area of approx 770 sq mile was searched by two observers and a recorder in a four-place helicopter provided by the USBR. Fissures were spotted from the air and checked on the ground. Locations were

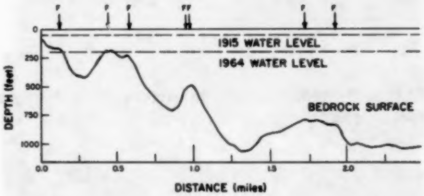


FIG. 6.—Bedrock Profile near Casa Grande, Ariz., Made by Gravity Survey; Arrows Show Fissure Locations

determined using the orthophotoquads as base maps. This procedure took three days and approx 15 h flying time (6). These studies led to a change in the proposed route of the Central Arizona Project (CAP) canal in the Picacho area between Phoenix and Tucson by the USBR, as the number of fissures and the amount of subsidence along a 20-mile segment of the canal were realized.

In the last 5 yr, studies of fissuring and subsidence have shown a relationship between fissure location and bedrock topography (2). Seismic and gravity profiles made across several subsiding areas in central Arizona have shown that bedrock "highs" or upward projections lie beneath many of the fissure locations (see Fig. 6). It is hoped that continued development of this type of research may lead to the ability to predict fissure locations in the near future. A single fissure was reported recently in the Paradise Valley area of Phoenix. Others are known in the area just east of Mesa. Their appearance in urban areas has not had a large impact as yet because they are still treated as a curiosity and because they have not yet caused any damage. Whether the ability to predict their occurence in or near urban areas in the future will be considered an advantage remains to be seen.

Conclusions

In the study of subsidence and earth fissures in central Arizona, aerial photography has been the indispensable tool. It has provided the means of recognizing a problem, a means by which the problem has been identified as a geologic hazard, and the means to monitor the expansion of both fissures and subsidence areas through updating data by periodic aerial photography, and by use of photographic base maps. The monitoring phase is becoming increasingly important as these geologic hazards move into the expanding urban areas.

Appendix.—References

1. Holzer, T. L., Davis, S. N., and Lofgren, B. E., "Faulting Caused by Groundwater Extraction in South-Central Arizona." *Journal of Geophysical Research,* Vol. 84, No. B-2, 1979, pp. 603–612.
2. Jachens, R. C., and Holzer, T. L., "Geophysical Investigations of Ground Failure Related to Ground-Water Withdrawal-Picacho Basin, Arizona," *Ground Water,* Vol. 17, No. 6, Nov.–Dec., 1979, pp. 574–585.
3. Laney, R. L., Raymond, R. H., and Winikka, C. C., "Maps Showing Water-level Declines, Land Subsidence, and Earth Fissures in South-Central Arizona," *Open-File Report 78-83,* United States Geological Survey, 1978.
4. Saxena, S. K. ed., *Evaluation and Prediction of Subsidence,* ASCE, Jan., 1978.
5. Schumann, H. H., and Poland, J. F., "Land Subsidence, Earth Fissures and Groundwater Withdrawal in South-Central Arizona, USA," First International Conference on Land Subsidence, Tokyo, Japan, 1968; also *Proceedings,* International Association of Hydrological Sciences, *Publication 88,* 1969, pp. 295–302.
6. Winikka, C. C., and Wold, P. D., "Land Subsidence in Central Arizona," *Proceedings,* Second International Symposium on Land Subsidence, Anaheim, Calif., 1977.

TRANSPORTATION
ENGINEERING JOURNAL

Designing for an Acceptable Wind Environment[a]

By Edward A. Arens,[1] M. ASCE

(Reviewed by the Aerospace Division)

Introduction

Tall or exposed buildings adjacent to public open spaces may cause local winds at ground level that are much more intense than winds found elsewhere at ground level. These winds may affect the comfort and safety of pedestrians and thus reduce the usefulness of the outdoor open spaces. In recent years, wind problems have become more common, as more tall buildings are built and as cities and building owners place increasing emphasis on public plazas and open space. Since both the cost and economic benefits of such plazas and open space may be very high, significant financial losses may occur when such spaces are rendered unusable due to wind.

The designers of buildings and their sites would benefit from being able to anticipate, in the planning stage, the possibility of local wind flow zones that cause unacceptable discomfort to users of outdoor space. If such zones are found, appropriate design decisions can eliminate them or direct pedestrians away from them.

This paper reviews present knowledge of pedestrian comfort in the wind and outlines how to design projects that avoid unacceptable wind environments.

Nature of Problem

Designers concerned with acceptable outdoor environments will have to address the following issues during the design of a building and its surrounding open space:

[a]Presented at the October 22–26, 1979, ASCE Annual Convention and Exposition, held at Atlanta, Ga.

[1]Assoc. Prof., Dept. of Architecture, Univ. of California, Berkeley, Calif. 94720.

Note.—Discussion open until August 1, 1981. To extend the closing date one month, a written request must be filed with the Manager of Technical and Professional Publications, ASCE. Manuscript was submitted for review for possible publication on November 5, 1979. This paper is part of the Transportation Engineering Journal of ASCE, Proceedings of the American Society of Civil Engineers, ©ASCE, Vol. 107, No. TE2, March, 1981. ISSN 0569-7891/81/0002-0127/$01.00.

1. How does the wind affect comfort and safety? What are the wind descriptors that measure comfort and safety? What are the limits of these descriptors at which the wind becomes uncomfortable or dangerous? These physiological and psychological questions can be answered by research in laboratories and in the field.

2. At any outdoor site, wind will probably exceed comfort or safety limits for a certain number of hours, minutes, or seconds during the year. The time during which the limits are exceeded, measured as a percentage of the total time that the project is occupied, indicate the users' probability of discomfort or danger. How large should these probabilities be? It is up to the designer and owner to decide acceptable probabilities for any given project, but certain levels have been suggested by researchers for use as design criteria.

3. How strong are the regional winds (as measured at the weather station) when local wind flows on site exceed comfort or safety limits? The relationship between the project's local wind and the wind of the surrounding region is a ratio based on the aerodynamic configuration of the project and its surroundings. The ratio varies with wind direction. Certain generalizations can be made, but complex configurations are best modeled in an appropriate wind tunnel.

4. How often do these excessive regional winds occur? This information, extracted from climatological records, determines the probability of the local wind on the site exceeding the limits for comfort or safety.

5. If local winds exceed the comfort or safety limits an unacceptable percentage of time, what design measures will reduce their strength so that the velocity limits are exceeded less often? These measures are devised by a combination of common sense, experience, and wind tunnel testing.

The following sections address these five issues, summarizing the state-of-the-art in acceptability criteria and design procedures.

Wind and Pedestrians

Wind influences comfort both mechanically, through its pressure effects and particle transport, and thermally, through wind chill. Pedestrian safety in the urban pedestrian environment is affected by mechanical pressure. Both mechanical and thermal effects are reviewed in the following to establish the wind speed limits above which comfort and safety are jeopardized. These limits form the basis for acceptability criteria used in design.

Mechanical Influences of Wind.—Wind influences comfort mechanically through pressure effects and particle transport. Wind pressure causes disturbance of clothing and hair, resistance to walking, and buffeting of the body and carried objects such as umbrellas. Comfort also is affected when the wind lifts dust and grit particles to eye level, or drives rain laterally into the eyes or beneath clothing. At higher velocities, wind interferes with walking and endangers people by causing them to lose their balance. At such velocities, eye damage from dust or possibly from flapping hair is also a safety problem. Some of the effects described are caused by the wind on a continuing basis, while others are precipitated by sudden unexpected peak gusts: a summary of such effects is provided in Refs. 1 and 18.

Wind at pedestrian level is accompanied by turbulence, and perceived as

varying velocity, gusts, or eddies. The intensity of turbulence for any given wind speed varies from place to place, tending to be greater in urban or built-up surroundings than in open countryside. The effects of turbulence on pedestrians may be described as an addition to mean velocity. An "equivalent steady wind" defined as giving the same comfort or safety effect as the turbulent wind was introduced by Hunt, Poulton, and Mumford (9):

$$u_s = \bar{u}(1 + a \cdot \mathrm{TI}) \qquad (1)$$

in which u_s = the equivalent steady wind; $\bar{u}$ = the mean wind speed; a = an empirically determined coefficient; and TI = the relative turbulence intensity (the root mean square of the instantaneous deviations from the mean velocity, divided by $\bar{u}$).

Hunt et al. found $a \simeq 3$ in experimental observations of the performance of pedestrians in a wind tunnel with controllable turbulence characteristics. The steady wind, u_s is a value greater than the mean, reflecting turbulent fluctuations superposed on the mean speed.

Jackson (12) used this relationship in defining a "standard equivalent mean wind speed" $\bar{u}_{se}$, in which terms he assembled a wide range of previously observed wind effects into a table combining the effects caused by steady uniform winds, turbulent wind fluctuations, and infrequently occurring peak gusts (see Table 1).

Certain standardized conditions were necessary to compile Table 1:

1. The standardized equivalent wind speed, $\bar{u}_{se}$, is measured at a height of 2 m over an averaging time of 5 min. Published observations for different averaging times are corrected to this time.
2. The horizontal relative turbulence intensity is 18%. Observations for which TI is unspecified are assumed to have occurred with a TI of 18%. Observations with specified TI values different from 18% are converted by the following relationship:

$$\bar{u}(1 + 3\,\mathrm{TI}) = \bar{u}_{se}(1 + 3 \times 0.18) \qquad (2)$$

3. For effects caused by a maximum gust of some minimum required duration, a value of $\bar{u}_{se}$ is calculated that would be expected to produce one such gust in the 5-min averaging period under 18% TI.

Wind effects observations as summarized in Table 1 and Ref. 17 are the basis for making judgments about wind acceptability in buildings and open space. Recommended acceptability criteria will be described later.

Thermal Influences of Wind.—The familiar concept of wind chill reflects the thermal influence of wind. In cool climates, wind increases the rate of cooling of the body by removing the insulating film of still air found next to skin and clothing in calm conditions. The increased rate of cooling may cause discomfort. In hot climates, the increased wind-induced convection and evaporation may be beneficial to comfort.

Thermal comfort is influenced by the following climatic variables: air, temperature, radiation (solar and terrestrial), humidity, and wind. The pedestrian's clothing and activity level are also important variables. Thermal comfort is a function

of body and skin temperature, the rate of heat transfer, and in overheated conditions, of skin wettedness as well.

To date, attempts to develop thermal models of human comfort outdoors have been confined to steady-state thermal balance models (1) in which thermal equilibrium is assumed to assure comfort. Such models have not been useful in practice primarily because thermal equilibrium with the surroundings requires 1 h–2 h continuous exposure to the outside environment. This is a rare situation for pedestrians, although such exposures may be experienced at bus stops and sports stadia. The average exposure is much shorter.

TABLE 1.—Wind Effects Versus Standard Equivalent Mean Wind Speed Under Standard Conditions*

Standard equivalent mean wind speed, in meters per second (1)	Effects observed or deduced (2)
0	Calm, no noticeable wind
2	Wind felt on face
4	Clothing flaps [5]
6	Newspaper reading becomes difficult (1)
8	Hair disarranged [5], dust and paper raised, rain
10	and sleet driven (1)
12	Control of walking begins to be impaired
14	Violent flapping of clothes [5], progress into
16	wind slightly slowed
18	Umbrella used with difficulty
20	Blown sideways [2], inconvenience felt walking
22	into wind, hair blown straight
	Difficult to walk steadily, appreciably slowed into wind [10]
	Noise on ears unpleasant
	Generally impedes progress
	Almost halted into wind, uncontrolled tottering downwind [10]
	Difficulty with balance in gusts [2]
	Unbalanced, grabbing at supports [2]
	People blown over in gusts [3]
	Cannot stand [3]

*The minimum gust duration required for each effect to be experienced is given, in seconds, in brackets [ ].

A computer program that iteratively calculates the thermal response of the body over a series of 1-min intervals shows promise for providing thermal design criteria for the wind environment (6). The effect of any period of exposure to outdoor climates on body temperature, heat transfer, thermal sensation, and comfort sensation can be predicted in this way. This model currently suffers from lack of experimental information on wind penetration or infiltration of clothing, and requires validation in outdoor conditions.

Design Criteria: Comfort, Safety, and Project Acceptability

With an understanding of wind effects on people, it is possible to suggest limits above which wind should not be permitted or is not desired. The Building Research Establishment, in England, recommended the following well-known values for mean wind at pedestrian height over an unspecified averaging period, in cool conditions (20): at 5 m/s, there is an onset of discomfort; at 10 m/s, it is definitely unpleasant; and at 20 m/s, it becomes dangerous.

Hunt et al. (9) refined these recommendations to the velocities given in Table 2. Murakami et al. (17) corroborated Hunt's limits, with some values somewhat more restrictive, finding control of walking difficult in the 10 m/s–15 m/s range. Nonuniformity of wind in time and space greatly affects walking and can cause the observed effect "walking difficult to control" in wind speeds as low as 3 m/s.

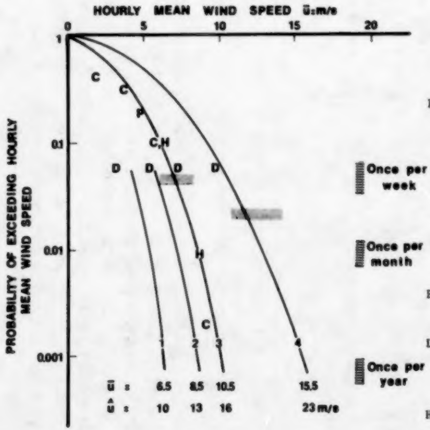
TABLE 2.—Wind Velocity Limits

Wind and effect (1)	Criterion[a] (2)
Steady uniform wind:	
For comfort and little effect on performance	$\bar{u} < 6$ m/s
For ease of walking	$\bar{u} < 13$ m/s–15 m/s
For safety of walking	$\bar{u} < 20$ m/s–30 m/s
Nonuniform winds ($\bar{u}$ varies by at least 70% over a distance less than 2 m):	
To avoid momentary loss of balance and to be able to walk straight	$\bar{u} < 9$ m/s
For safety (for elderly people this criterion may be too high)	$\bar{u} < 13$ m/s–20 m/s
Gusty winds[b]:	
For comfort and little effect on performance	$u_s < 6$ m/s
Most performance unaffected	$u_s < 9$ m/s
Control of walking	$u_s < 15$ m/s
Safety of walking	$u_s < 20$ m/s

[a] $\bar{u}$ is averaged over short periods, on the order of seconds.
[b] u_s is defined in Eq. 1.

With a wind speed limit in hand, the designer must judge what percentage of the time (or with what frequency in a year or season) it may be exceeded. The percentage of time exceeded equals the probability of discomfort or danger. For comfort limits, 10%–20% may seem reasonable. For safety limits, lower percentages of time exceeded (say 0.1%) should be applied to the higher velocities. Various researchers have suggested limits and acceptable frequencies for the limits. These have been summarized and compared by Melbourne (15).

Fig. 1 is a reproduction of a figure from Ref. 15, with some additions as noted later. The figure compares the various criteria using the probability of exceeding various hourly mean speed limits in any given year. The probabilities are adjusted to apply to half the hours in the year representing the daylight hours when most buildings are occupied. The curves imply a relative turbulence intensity of 15%.



Legend for Comfort Criteria Graph

CRITERIA		SYMBOL
MELBOURNE (16)		
Acceptable below for:		
Walking		1
Stationary, short exposure		2
Stationary, long exposure		3
Unacceptable, dangerous above:		4
DAVENPORT (5)		
Acceptable for:		
Walking fast	If P[$\bar{u}$ > 10] < 0.05	D
Strolling	If P[$\bar{u}$ > 7.5] < 0.05	D
Standing, Sitting, Short Exposure	If P[$\bar{u}$ > 5.5] < 0.05	D
Standing, Sitting Long Exposure	If P[$\bar{u}$ > 3.5] < 0.05	D
PENWARDEN AND WISE		
Acceptable	If P[$\bar{u}$ > 5] < 0.2	P
LAWSON (13)		
Acceptable	If P[$\bar{u}$ > 6 to 8] < 0.04	
Unacceptable	If P[$\bar{u}$ > 11 to 14] < 0.02	
HUNT, POULTON & MUMFORD (9)		
Acceptable for Strolling	If P[$\bar{u}$ > 6] < 0.1	H
Acceptable for Walking	If P[$\bar{u}$ > 9] < 0.01	H
COHEN, ET AL (3)		
Data from Table 3.		C

FIG. 1.—Comparison of Various Criteria for Environmental Wind Conditions for Daylight Hours, for Relative Turbulence Intensity of 15%

Isyumov (10,11) presents a plot of comfort and safety versus hourly mean wind speed similar to Melbourne's but expressed in terms of number of occurrences per year where the hourly mean wind speed is greater than the comfort limits indicated. The return periods are calculated using a method by Davenport (4).

Cohen et al. (3) recommends a number of acceptability criteria based on their extensive field observations (see Table 3). These findings have been incorporated into Fig. 1 correcting for daylight hours for consistency with Melbourne's method. The safety-related values show lower wind speed limits in the acceptability criteria than the other researchers, almost certainly because the peak gusts recorded per hour in this investigation tended to three times the hourly mean, whereas such gusts in Melbourne's assumed turbulence would

TABLE 3.—Pedestrian Safety/Comfort Standards for Urban Winds

Activity area (1)	Hourly mean wind speed, in meters per second (miles per hour) (2)	Permitted occurrence frequency, as a percentage (3)	Permitted frequency considering only daylight hours, as a percentage (4)
All pedestrian areas—limit for safety	9.1 (20)	0.1 (≈10 h/yr)	0.2
Major walkways, especially principal egress path for high-rise buildings	9.1 (20)	0.1	0.2
Other pedestrian walkways, including street and arcade shopping areas	6.4 (14)	5	10
Open plazas and park areas walking, strolling activities	3.6 (8)	15	30
Open plaza and park sitting areas, open-air restaurants	2.3 (5)	20	40

be only 1.5 times the hourly mean. Because the turbulence intensity was not reported with Cohen's data, it is not possible to adjust these values to make them fully comparable. The comparison does show the significance of assumed turbulence intensity in these acceptability criteria for safety purposes, where the influence of the peak gust predominates.

Penwarden (18) analyzed cases of shopping centers that had experienced wind complaints. He found that in centers where the single limit of 5 m/s hourly mean wind speed was exceeded 20% of the time or more, the owners invariably spent money to add protective screens or roofs. Centers with frequencies of 10%–20% caused complaints but no remedial action was taken. Few complaints were registered at centers with frequencies below 10%. Penwarden's study gives the most concrete economic evidence in support of specific acceptability criteria to date.

Caution should be exercised in applying such criteria to wind data available from the United States National Oceanic and Atmospheric Administration (NOAA). In a nonsteady wind, the length of the averaging interval will affect the value of $\bar{u}$ associated with the maximum wind effect observed during that interval. This is because most wind effects are sensed by the pedestrian over very short intervals of time, 1 sec–10 sec, while wind is usually observed and recorded for longer intervals, of 1 min or more. Melbourne's and Cohen's effects on the chart are either observed against hourly menas, or converted to be so. The actual effects, however, may refer to instances that may have occurred only once, during a gust, within that hour.

The NOAA wind records available in the United States are termed hourly mean wind speeds, but are actually 1-min means measured once an hour. This means that the distribution of windspeeds from NOAA data will show greater

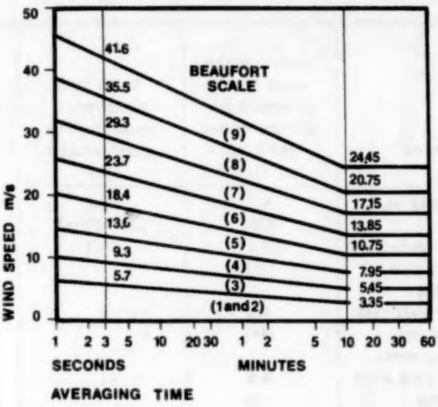


FIG. 2.—Variation of Wind Speed with Averaging Time, for Velocity Ranges of Beaufort Scale

variability than would true hourly means. The pedestrian effects specified for gusts implied within Melbourne's hourly mean wind speeds would therefore occur at higher mean wind speeds when using NOAA data, speeds closer to those of the gust itself. Fig. 2, reproduced from Lawson (14), gives an example of the variation of wind speed with averaging time at a relative turbulence intensity of 0.28. One may estimate from this that the hourly mean criteria should be 25% higher when using NOAA data; that the 5-min mean wind speed limits described above should be 20% higher; and that criteria based on direct peak gust measurement (2 sec–3 sec), should be 25% lower.

Finally, caution is needed when using any of the criteria based on mechanically-caused wind discomfort, for thermal discomfort in cool environments usually begins at lower velocities than mechanical discomfort. Reliance on these may underestimate the actual discomfort, especially for prolonged or relatively inactive outdoor activities.

Determining Surface Winds in Proposed Built Environment

The designer should consider outdoor comfort and safety early in the design process. Because the relationships between the physical form of a building and site, the climate, and resulting comfort and safety around it are complex, he or she may have to follow a climatic design process in order to find a satisfactory solution. The process, basically: (1) Determines the climatic characteristics of the site and the preliminary project, partly by model tests; (2) assesses its effects on the acceptability of the project; and (3) modifies the project design and tests the climate until a solution is reached.

Winds at pedestrian level often will be strongly affected by the building, planting, and grading configuration of the project. Briefly, there are basically three flow fields that cause enhanced wind at pedestrian level: (1) Vortices at the base of the windward face of a building caused by greater pressures on that face at higher elevations; (2) flows caused by pressure differences between low pressure regions at the sides and lee of a building and the relatively higher pressure regions at the windward side (open passageways between these regions will experience greatly enhanced winds); and (3) flows through constrictions between buildings. Two publications by Gandemer (7,8) provide visualizations of wind flow around a wide variety of building configurations. They are very useful for obtaining an intuitive feel for wind behavior around buildings.

Penwarden and Wise (19) have provided generalized rooftop-to-ground-level velocity ratios for such configurations that have been widely quoted. However, note that these ratios do not apply when buildings of similar height are located in the upwind direction.

If the configuration of a yet-unbuilt project seems likely to cause high local winds, it should be tested in model form in a wind tunnel. This technique is also useful for defining winds on existing sites, since the flow strength and direction can be controlled during the tests. Physical modeling with limited field verification is most desirable.

Physical modeling requires the use of a specialized wind tunnel that reproduces the boundary-layer conditions above the actual site. Both the velocity and the turbulence intensity profiles should be modeled to scale. The most satisfactory means of achieving this at present is to generate the boundary layer with turbulence generators and long fetches of roughness similar to that of the terrain upstream of the project site. The testing of architectural models for environmental wind conditions may be carried out at low wind speeds (5 m/s–10 m/s) because turbulent flow patterns around bluff (sharp-edged) objects do not vary over a wide range of velocities. This is fortunate, for it allows tests to be performed at costs comparable to other design consulting fees.

Velocities measured in the wind tunnel are nondimensionalized and are expressed as a percentage of a reference velocity. The reference velocity in the tunnel is measured at a reference height, often chosen as the height (at model scale) of the wind instrumentation of the weather station providing climatological wind data. By relating wind tunnel measurements to climatological data, wind speed frequency distributions are found for important locations within and around the proposed project.

Measurements are normally made at a network of grid points on the project model. A hot-wire anemometer is used to measure wind speed and turbulence

at each grid point. Its small size allows it to measure speeds within distances of the order of millimeters above the surface, representing pedestrian height at model scale. In addition, when the wire is held vertically, the hot-wire anemometer is insensitive to the azimuth angle of approaching wind (within a sector of about 270°). Since the wind turbulence at pedestrian level near buildings often results in rapid lateral directional changes, it is important that the measuring instrument exhibit minimal directional sensitivity in the horizontal plane.

The entire network of points on the model is tested separately for each wind direction, the number of directions normally corresponding to the number of points of the compass considered in the meteorological wind data base.

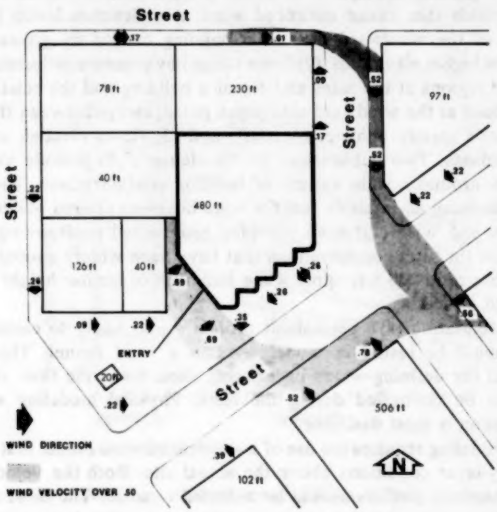


FIG. 3.—Street-Level Wind Directions and Velocities during West Wind, High-Rise Project

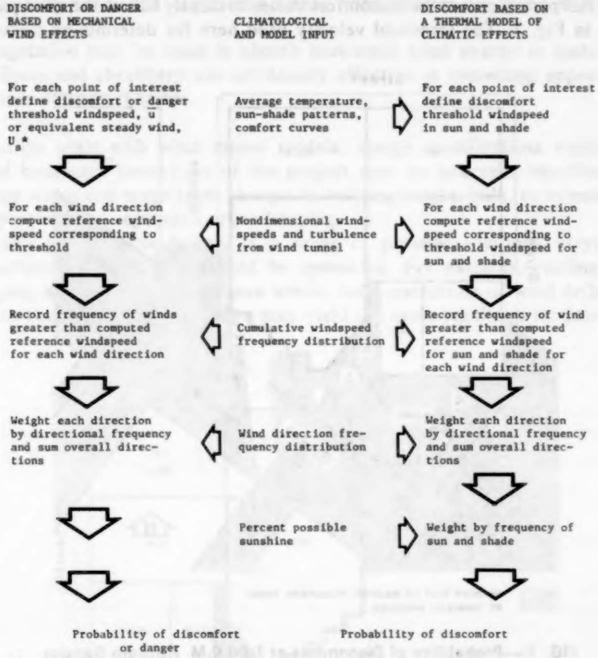
Fig. 3 is a representation of a strong air flow pattern occurring around a wind tunnel model of a San Francisco high-rise project during exposure to west winds. The building's entrance is located on the southwest side of the building. The wind directions at several grid points are marked and the velocities expressed as decimal fractions of the velocity at the reference height, which in this case is 132 ft, the height of the anemometer at the local weather station in the city.

Determining Frequencies of Uncomfortable or Dangerous Winds in Proposed Built Environment

Wind data are usually recorded at airports in open terrain. For most sites

wind information must be extrapolated geographically from the recording station to the vicinity of the site. The effects of topography, vegetation, and structures must be carefully considered in this extrapolation.

The meteorological data base should provide information necessary to determine the amount of time that pedestrians will be uncomfortable or endangered on the site. This requires predictions of expected winds preferably by time of day, and if thermal comfort is being estimated, coincident data on temperature and sun. The most useful data format is a cumulative directional frequency



*See Table 2

FIG. 4.—Process for Determining Probabilities of Discomfort or Danger

distribution providing the percentage of time that each wind velocity is exceeded for each wind direction. The National Climatic Center of the NOAA may have this information in existing published summaries, although most locations have not been analyzed in detail. The Center will prepare such summaries from original hourly observations upon request. It also provides magnetic tapes with the hourly observations, usually with 10 yr/tape. The data from the tape is then commonly fitted to a model such as the Weibull distribution to provide a smooth frequency distribution.

The frequency of discomfort may then be determined for either thermal or mechanical wind effects by following the procedure in Fig. 4 taken from Ref. 1. The same approach may be used to predict dangerous wind velocity frequencies. The procedure for thermal comfort is carried out for selected hours of the day to incorporate sun/shade patterns and temperature data into the comfort determination.

Fig. 5 shows the probability of discomfort around the high-rise project considering all wind conditions expected in San Francisco in the early afternoon during the autumn season. The predominance of west winds over other winds during this period causes the discomfort values to closely follow the wind patterns shown in Fig. 3. The threshold velocity used here for determining discomfort

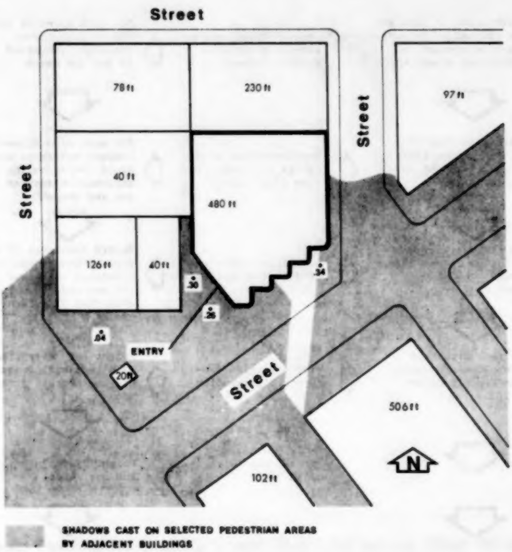


FIG. 5.—Probability of Discomfort at 1:00 P.M. Autumn Season

is 5 m/s. The shadow patterns for this period are also shown, to give a suggestion of the thermal environment.

Avoiding and Mitigating Wind Problems

If the expected comfort levels are unacceptable, design modifications to improve conditions may have to be investigated. The following measures are appropriate for tall buildings (2,3):

1. Large slab buildings should not be oriented in a direction normal to the prevailing wind to avoid downwash on the windward face. Circular and polygonal

towers tend to have advantageous wind climates at ground level because of reduced downwash.

2. Tall buildings benefit from significant horizontal projections to break up downward-directional winds. Protective awnings and canopies near ground level need to be large to influence wind over appreciable areas around buildings.

3. Important pedestrian thoroughfares and building entrances should not be planned at the windward corners of tall slab buildings as these are regions of accelerated corner flows.

4. Openings through buildings near the ground, especially with openings facing the prevailing wind, will experience strong winds unless revolving doors are used.

5. Vegetation may be used to absorb horizontal wind energy in pedestrian areas. Trees and shrubbery are not usually effective at protecting appreciable areas from downdraft winds.

Based on trials with wind tunnel models, design modifications within the site and budgetary limitations of the project may be selected. Modifications to reduce winds can range from changes in building height, bulk, or orientation, to the provision of vegetation or landscaping.

For any given problem there is a range of possible solutions varying in effectiveness and cost that should be optimized. For example, roofing over a shopping mall would surely reduce winds, but construction of wind deflectors or latticework or use of vegetation may yield the same results at far less cost.

FIG. 6.—Percentage Reductions in Wind Speeds Near Entry due to Proposed Canopy above Entry

The influence of such devices on sunlight also should be considered.

A model of the modified design is tested in the wind tunnel and the analysis of comfort repeated. Further modifications are again suggested and the process is repeated until a satisfactory design is obtained.

Fig. 6 shows the reduction in the wind speeds at the sidewalk for the west wind case when a canopy is extended over the entrance. The reduction in expected discomfort on the sidewalk may be determined from this, allowing the designer to assess the feasibility of the canopy.

Conclusions

Wind effects on comfort and safety have been described and limits for acceptable velocities proposed by various authors have been summarized. These limits are strongly influenced by the averaging interval selected, and by the turbulence component of the wind. The paper presents suggested design criteria for the amount of time that these velocity limits may be exceeded.

The designer of a project affecting pedestrians outdoors may estimate whether the project meets these criteria by synthesizing information on the aerodynamic characteristics of the project and climatological information on wind frequency distributions for the region. Some general problem-causing building geometries can be identified, but many urban sites are sufficiently complex to justify model testing in a wind tunnel. A procedure for determining project acceptability through model testing is outlined.

Acknowledgments

The writer wishes to thank Steven Margulis and Emil Simiu for their thoughtful reviews, and Susan Johnson and Ana Salazar for help with the manuscript. This paper was written at the Center for Building Technology, National Bureau of Standards, Washington, D.C.

Appendix.—References

1. Arens, E. A., and Ballanti, D. B., "Outdoor Comfort of Pedestrians in Cities," *Proceedings of the Conference on the Metropolitan Physical Environment*; also *General Technical Report NE-25*, United States Department of Agriculture Forest Service, 1977, pp. 115–129.
2. Aynsley, R. M., "Effects of Airflow on Human Comfort," *Building Science*, Vol. 9, 1974, pp. 91–94.
3. Cohen, H., McLaren, T., Moss, S., Petyk, R., and Zube, E., "Pedestrians and Wind in the Urban Environment," *UMASS/IME/R-77/13*, University of Massachusetts, Amherst, Mass., 1977, 124 p.
4. Davenport, A. G., "On the Statistical Prediction of Structural Performance in the Wind Environment," presented at the April 1971, ASCE National Structural Engineering Meeting, held at Baltimore, Md. (Preprint 1420).
5. Davenport, A. G., "An Approach to Human Comfort Criteria for Environmental Wind Conditions," CIB/WMO Colloquia, Teaching the Teachers, Swedish National Building Research Institute, Stockholm, Sweden, 1972.
6. Gagge, A. P., Nishi, Y., and Nevins, R. G., "The Role of Clothing in Meeting FEA Energy Conservation Guidelines," *Transactions*, American Society of Heating Refrigerating, and Air-Conditioning Engineers, Vol. 82, 1977, pp. 234–247.
7. Gandemer, J., "Discomfort Due to Wind Near Buildings: Aerodynamic Concepts,"

report of the Centre Scientific et Technique du Batiment, Paris, France, translated and issued as *NBS Technical Note 710-9*, National Bureau of Standards, Washington, D.C., 1978, 41 p.

8. Gandemer, J., and Guyot, A., "Intégration due Phénoméne Vent dans la Conception du Milieu Bati," report of the Secrétariat Général du Group Central des Villes Nouvelles, Paris, France, 1976, 130 p.
9. Hunt, J. C. R., Poulton, E. C., and Mumford, J. C., "The Effects of Wind on People: New Criteria Based on Wind Tunnel Experiments," *Building and Environment*, Vol. 11, 1976, pp. 15–28.
10. Isyumov, N., "Studies of the Pedestrian Level Wind Environment at the Boundary Layer Wind Tunnel Laboratory of the University of Western Ontario," *Journal of Industrial Aerodynamics*, Vol. 3, 1978, pp. 187–200.
11. Isyumov, N., and Davenport, A. G., "The Ground Level Wind Environment in Built-up Areas," *Proceedings Fourth International Conference on Wind Effects on Buildings and Structures*, Cambridge University Press, Cambridge, England, 1977.
12. Jackson, P. S., "The Evaluation of Windy Environments," *Building and Environment*, Vol. 13, 1978, pp. 251–260.
13. Lawson, T. V., "The Wind Environment of Buildings: A Logical Approach to the Establishment of Criteria," *Report No. TVL 7321*, Dept. of Aeronautical Engineering, University of Bristol, Bristol, England, 1973.
14. Lawson, T. V., "The Wind Content of the Built Environment," *Journal of Industrial Aerodynamics*, Vol. 3, 1978, pp. 93–105.
15. Melbourne, W. H., "Criteria for Environmental Wind Conditions," *Journal of Industrial Aerodynamics*, Vol. 3, 1978, pp. 241–249.
16. Melbourne, W. H., and Joubert, P. N., "Problems of Wind Flow at the Base of Tall Buildings," *Proceedings of the Third International Conference on Wind Effects on Buildings and Structures*, Saikon Co., Tokyo, Japan, 1971.
17. Murakami, S., Uehara, K., and Deguchi, K., "Wind Effects on Pedestrians: New Criteria Based on Outdoor Observation of over 2000 Persons," *Proceedings of the Fifth International Conference on Wind Engineering*, Pergamon Press, New York, N.Y., Vols. 1 and 2, 1979 and 1980.
18. Penwarden, A. D., "Acceptable Wind Speeds in Towns," *Building Science*, Vol. 8, 1973, pp. 259–267.
19. Penwarden, A. D., and Wise, A. F. E., "Wind Environment Around Buildings," Building Research Establishment Report, London, England, HMSO, 1975.
20. Wise, A. F. E., "Wind Effects Due to Groups of Buildings," *Current Paper 23/70*, Building Research Station, Watford, England, 1970, 16 p.

[illegible]

TRANSPORTATION ENGINEERING JOURNAL

Downgrade Speed Characteristics of Heavy Vehicles

By Abishai Polus,[1] M. ASCE, Joseph Craus,[2] and Izhak Grinberg[2]

(Reviewed by the Highway Division)

Introduction

The speed performance of heavy vehicles on downgrades is needed for capacity calculations, for determination of lane requirements, and for analyses such as user costs in economic studies. Such performance information, however, is rarely available in published form.

This study, which is based on field observations, analyzed the speed characteristics of heavy vehicles—trucks and buses—on selected rural downgrades throughout Israel (2). An attempt was then made to model the approach-speed gradient, which is defined as the ratio of the difference between the approach (level) speed and the (lower) speed on the grade to the approach speed, in terms of the length and slope of the grade. Finally, based on a previously developed model (2) for determining passenger-car equivalencies (P.C.E.), several values were calculated and presented for the downgrades under investigation.

Background of Study

Very little information is available on the operational characteristics of heavy vehicles on a downgrade, although their performance limitations are widely accepted. While upgrade performance is primarily influenced by engine capabilities, or more specifically weight-to-horsepower ratio, downgrade performance is affected by complexity of components, which may include the length and

[1] Sr. Lect. of Civ. Engrg., Dept. of Civ. Engrg., Technion—Israel Inst. of Tech., Haifa, Israel; currently Visiting Assoc. Prof., Dept. of Civ. Engrg., West Virginia Univ., Morgantown, W. Va. 26506.

[2] Prof. of Civ. Engrg. and Head of Transportation Research Inst., Technion—Israel Inst. of Tech., Haifa, Israel.

[3] Research Engr., Dept. of Civ. Engrg., Technion—Israel Inst. of Tech., Haifa, Israel.

Note.—Discussion open until August 1, 1981. To extend the closing date one month, a written request must be filed with the Manager of Technical and Professional Publications, ASCE. Manuscript was submitted for review for possible publication on November 20, 1979. This paper is part of the Transportation Engineering Journal of ASCE, Proceedings of the American Society of Civil Engineers, ©ASCE, Vol. 107, No. TE2, March, 1981. ISSN 0569-7891/81/0002-0143/$01.00.

grade of slope being traversed as well as of the previous and following slopes, sight distance, and driver skill and attitude. It is difficult, therefore, to establish any definite relationships between speed and rate of grade for heavy vehicles on downgrades.

The American policy on geometric design of rural highways (1) suggests that compared to level operation, heavy vehicles on downgrades show an increase in speed for grades up to about 5% and a decrease in speed for grades of about 7% or steeper. The Australian policy (6) states that commercial vehicles going downhill often travel at speeds greater than their usual level running speed, provided the grade does not exceed 6%.

One of the few studies which deals with downhill speeds of heavy vehicles, in this case trucks, was performed by the State of California, Department of Public Works (7). Field-speed observations were made on several grades throughout the state although no distinction was made between loaded and empty trucks. The length of grade ranged between 1.5 miles and 6.2 miles and the slope ranged between 2% and 7%. The results of these observations did not produce a set of standard curves from which downhill speeds could be obtained for use in any situation; however, they appeared to show a distinct difference in the behavior of trucks on long downgrades as compared to shorter downgrades. for long grades, trucks were observed to slow down to a "crawl" speed and maintain that speed until not far from the bottom of the grade. For short, steep grades, their speeds were slower near the summit but increased uniformly down the grade.

Another study, by the Midwest Research Institute (5) analyzed downgrade performance utilizing the simulation technique and suggested a measure for the rating of grades. It found that both length and rate of grade influenced truck performance on downgrades, the percent of grade factor being more significant.

Data Collection

Speed and weight field observations were performed at six rural downgrades, the characteristics of which are presented in Table 1. A detailed site study

TABLE 1.—Downgrade Sites and Pertinent Characteristics

Site number (1)	Site name (2)	Length of down-grade, in meters (3)	Average slope, in percent (4)	Number of measurement points (5)	Number of Vehicles: Trucks (6)	Number of Vehicles: Buses (7)	Number of Vehicles: Passenger cars (8)
1	Carmiel 1	500	9.0	5	69	12	25
2	Carmiel 2	530	4.5	3	118	24	42
3	Bat-Shlomo	500	4.8	3	97	2	54
4	Lavie	850	3.5	3	83	19	61
5	Elifelet	950	6.8	3	56	12	59
6	Migdal-Ha'emek	1,150	7.7	6	56	13	52

led to the selection of only two-lane rural downgrades with a tangent section, where the approach stretch was about level. Another consideration in the choice of sites was that they have the greatest possible resemblance to one other in terms of design speed, width of pavement, and width of shoulder.

The speed data was gathered with a portable data collection system which was composed of three major components: (1) Tapeswitches; (2) video-tape recording and monitoring equipment; and (3) a digital clock. The array of 8 ft long tapeswitches was taped to the road in pairs to transmit electrical pulses upon the detection of vehicle presence. The pairs of tapeswitches were taped at a given constant distance apart and connected by telephone cable to the main control box. The time was recorded to 1/100 of a second accuracy upon activation of a pair of switches. The speed of a vehicle between the first and last pair of tapeswitches was transformed to a profile of spot speeds along the grade for as many points as the number of tapeswitch pairs used. The speeds were gathered at three or six cross sections along each of the six downgrades studied. The first cross section was on the approach to the grade; the other points were about evenly distributed along the grade. During the speed data collection, information about the class and load of the vehicle was also collected. Speed profiles were later devised for each truck load group separately, as well as for buses and passenger vehicles. In order to obtain the necessary information and ranges of weight classifications, a study of static weight was performed on two sites utilizing a British made scale.

Speed Characteristics

Speed measurements were recorded separately for the following classes of heavy vehicles: (1) Loaded trucks; (2) empty trucks; (3) buses; and (4) passenger vehicles. The series of spot speed reading for each vehicle was transformed to a weighted space speed along the grade, based on the distance between successive pairs of tapeswitches. Each such speed reading was assumed to represent the space speed before and after the particular tapeswitch for up to half the distance of the preceding and following speed. A typical example

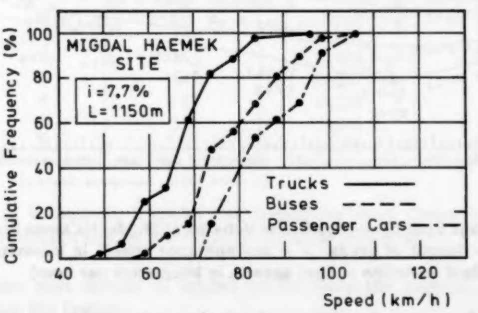


FIG. 1.—Cumulative Distribution Function of Space Speed along Migdal-Ha'emek Downgrade (i = percent of grade; and L = length of grade)

of the cumulative distribution function of downgrade space speeds at the Migdal-Ha'emek site (Site #6) is presented in Fig. 1. Passenger cars can be

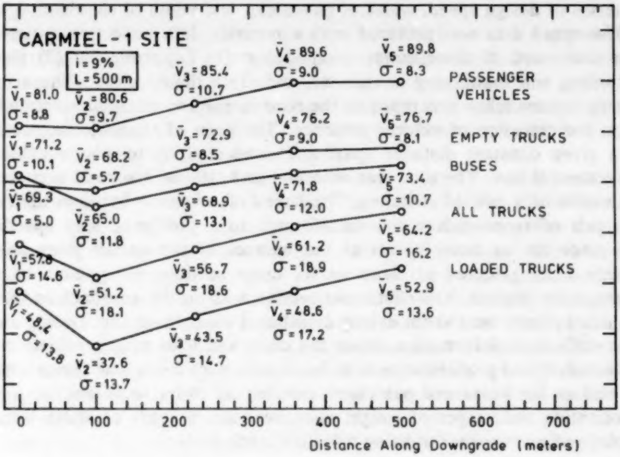


FIG. 2.—Average Spot Speed Profiles of Vehicles at Carmiel 1 Site (i = percent of grade; L = length of grade; $\bar{V}$ = average spot speed, in kilometers per hour; and σ = standard deviation of spot speeds, in kilometers per hour)

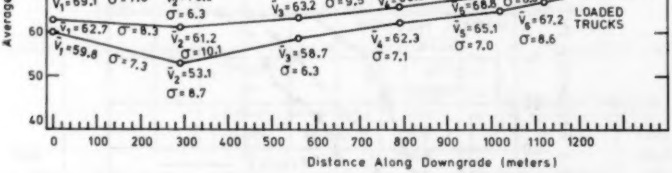


FIG. 3.—Average Spot Speed Profiles of Vehicles at Migdal Ha'emek Site (i = percent of grade; L = length of grade; $\bar{V}$ = average spot speed, in kilometers per hour; and σ = standard deviation of spot speeds, in kilometers per hour)

seen to have faster speeds while trucks, both empty and loaded, travel at lower speeds. It was further found, at this site and at others, that the speed distribution of buses is closer to that of cars than to that of trucks. This finding may

be explained by the lower weight-to-horse power ratio of buses and their usually better braking systems, which caused bus drivers to feel safer at higher speeds and to develop faster downgrade speeds.

A profile of average spot speeds along the downgrade was performed and Fig. 2 and Fig. 3 present the findings which include the standard deviation

TABLE 2.—Average Spot Speed, $\bar{V}$, and Standard Deviation, σ, according to Number of Axles of Trucks, in kilometers per hour (Carmiel 1 Site)

Variable (1)	Axle configuration and number of trucks (2)	Cross Section 1 — (3)	2 100 m (4)	3 230 m (5)	4 360 m (6)	5 500 m (7)
		Distance along grade				
$\bar{V}$	2 axles	61.2	56.8	60.8	65.7	67.8
σ	37 trucks	11.1	13.5	14.1	14.2	12.2
$\bar{V}$	3 axles	55.1	48.7	55.2	58.6	63.5
σ	17 trucks	17.7	22.5	23.7	23.7	20.0
$\bar{V}$	4 axles	54.6	44.3	50.3	56.0	57.6
σ	10 trucks	13.5	15.3	19.7	21.0	17.9
$\bar{V}$	5 axles	42.1	31.4	39.4	41.7	47.8
σ	5 trucks	22.1	19.6	16.3	18.0	14.8

TABLE 3.—Weight of Empty and Loaded Trucks according to Number of Axles

Type of truck (1)	Number of axles (2)	Number of trucks (3)	Average weight, in kilograms (4)	Standard deviation of average weight, in kilograms (5)
Loaded	2	49	17,931	2,230
Empty	2	62	9,210	1,677
Loaded	3	15	27,460	3,344
Empty	3	15	16,819	1,973
Loaded	4	14	35,256	3,405
Empty	4	8	18,220	952
Loaded	5	12	50,905	7,492
Empty	5	7	26,963	2,284

for each class of vehicle for Carmiel and Migdal-Ha'emek (Sites #2 and #6), respectively. Prominent among the findings were the following:

1. The average spot speeds of loaded trucks were the slowest, and those of passenger cars the fastest.

2. The average speeds at the end (bottom) of each downgrade were significantly higher than the initial speeds. This trend was more pronounced for passenger

cars and buses, and less significant for loaded trucks although there was some increase speed for all groups.

3. For trucks, average speeds at the second measuring station—whose distance from the beginning of the grade was about 20%–30% of the total site length—were the lowest in all cases. This finding was especially noticeable for loaded trucks. Beyond the second point, the speed increased rather uniformly, owing, perhaps, to the relatively short length of the grade. No manifestation of sustained speed was detected. It is reasonable to assume however that long grades induce a sustained downgrade "crawl" speed, as mentioned in a previous study (7).

4. There appears to be a distinct difference in the behavior of loaded trucks and buses; the flow characteristics of the latter on downgrades are similar to passenger car characteristics.

A further breakdown of the speed data was made according to truck axle configuration and number. Results for the five cross sections at Carmiel site are presented in Table 2. A significant reduction in average spot speed may be observed for an increasing number of truck axles. Table 3 presents the relationship between number of axles and weight of empty and of loaded trucks, as obtained prior to the main study in a pilot study especially designed for the collection of weight data. It can be seen that the average weight of loaded trucks is almost twice that of empty trucks.

Approach-Speed Gradient

During the course of this study, several measures, in two categories, were developed: (1) Measures based on acceleration or deceleration; and (2) measures based on spot speeds. The acceleration or deceleration measures, considered initially, were eliminated at an early stage, since the portable data collection system was not designed for this type of data, which had to be derived indirectly. It was decided, therefore, to concentrate on measures based on spot speeds by comparing the spot speed at each station, i.e., location of a pair of tapeswitches, to the approach speed, i.e., the level operation just before the downgrade. For the downgrades studied, it was found that the spot speed at the second point along the grade was the lowest among all stations. A measure of the approach-speed gradient was defined, and is given by Eq. 1 as:

$$GV_k = \frac{100\,(V_{k1} - V_{k2})}{V_{k1}} \quad \text{. . . . . (1)}$$

in which GV_k = approach-speed gradient for vehicle k, in percent; V_{k1} = spot speed for vehicle k at section 1 (approach speed); and V_{k2} = spot speed for vehicle k at section 2.

The average of all approach-speed gradients was calculated, at each site, for all classes of vehicles. This is given by Equation 2:

$$\overline{GV} = \frac{1}{n}\sum_{k=1}^{n} GV_k \quad \text{. . . . . (2)}$$

in which $\overline{GV}$ = mean approach-speed gradient for a certain class of vehicles,

in percent; GV_k = approach-speed gradient for vehicle k; and n = number of vehicles of a given class at one site.

The values of the mean approach-speed gradient were calculated for loaded trucks, empty trucks, buses, and passenger vehicles and are presented in Table 4. The greatest reduction in speed was observed for loaded trucks; buses had the greatest relative increase in speed, probably because of improved braking systems and the greater skill of their drivers, compared with passenger car drivers. By plotting the mean gradient against the slope of the downgrade, an increase in slope was found to be directly related to a gradient increase, a tendency that was much more pronounced for slopes of 6% or more. Greater lengths were also found to contribute to an increase in approach-speed gradient, although their contribution was relatively less compared to the slope variable. These associations are shown in Table 4.

A basic hypothesis thus guided the calibration of models during the research; the mean gradient increases as the length and average grade increase.

TABLE 4.—Mean Approach-Speed Gradient for Four Vehicle Classes

Site number (1)	Site (2)	Average slope, in percent (3)	Length, in meters (4)	Mean Approach-Speed Gradient, in percent[a] Loaded trucks (5)	Empty trucks (6)	Buses (7)	Passenger cars (8)
1	Carmiel 1	9.0	500	19.9	2.0	9.1	0.5
2	Carmiel 2	4.5	530	0.2	−4.0	−6.7	−2.0
3	Bat-Shlomo	4.8	500	5.6	−7.1	—	−5.7
4	Lavie	3.5	850	2.8	−5.9	−6.5	−2.6
5	Elifelet	6.8	950	3.7	−5.2	−8.1	−5.1
6	Migdal Ha'emek	7.7	1,150	11.1	−2.0	−4.3	−3.0

[a]Minus value represents an increase in mean speed; plus value, a decrease in mean speed.

Utilizing the regression techniques, a few simple models were constructed, having the following general form:

$$\overline{GV} = A\,[f_1(i)f_2(L)] + B \qquad (3)$$

in which $\overline{GV}$ = mean approach-speed gradient for a certain class of vehicles, defined by Eq. 2; i = average slope of downgrade, in percent; L = length of downgrade, in meters; f_1, f_2 = functions; and A, B = calibration constants.

The model that had the best fit to the data had an exponential nature, and its calibration is presented by Eq. 4:

$$\overline{GV} = 0.0046\,Le^{i} + 0.4; \quad R^2 = 0.93, \ F = 80.0 \qquad (4)$$

in which e = the base of the natural logarithm, and all other parameters are defined in Eq. 3. The coefficient of determination R and F statistics are defined in the literature (4). The graphical presentation of this equation is presented in Fig. 4, along with the measured data. It can be seen that mean approach-speed

gradients have a limiting value of 0.4%, which may be interpreted as the "natural gradient." For values of Le^i that are less than 130, the mean gradient is less than 1%. Beyond 130, the mean gradient increases sharply and equals 5% for a value of 1,000. The contribution of the average slope of downgrade, *i*, has an exponential nature, and its contribution to the average gradient is more significant than is that of the length of the downgrade. After the exponential

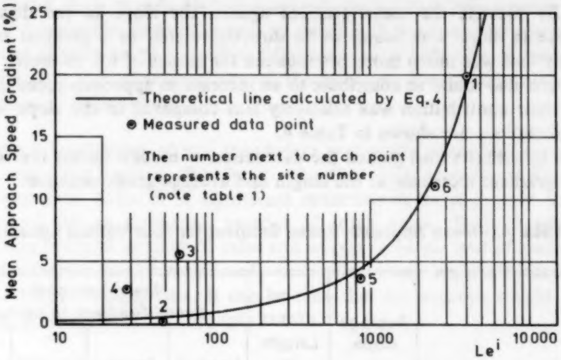


FIG. 4.—Suggested Model of Mean Approach-Speed Gradient as Related to Length and Slope of Downgrade

nature of the relationship was established, truck passenger-car equivalents were determined.

Passenger Car Equivalencies

An earlier section of this study dealt with the detailed development of a revised method for the determination of passenger-car equivalencies (PCE) (2). The PCE was evaluated by the ratio of the average delay caused by one truck to the average delay caused by one passenger car. This proposed method is rather different from the regular approach (3), which is based on the ratio of the theoretical number of passings of one truck to the average theoretical number of passings of one passenger car.

Based on the cumulative distribution of spot speeds of passenger cars, a typical example of which is presented in Fig. 1, and on the calculated average speed of empty trucks, loaded trucks, and buses at all six sites, passenger car equivalents have been calculated utilizing the delay measure (2). For comparison purposes, the PCE values were also calculated by the proposed method for a level stretch of two-lane highway. The results are presented in Table 5. This table shows that considerable differences in passenger car equivalents exists between empty and loaded trucks, as would be expected, owing to the difference in speeds between the two classes. Also, buses were found to have relatively low passenger-car equivalents on most downgrades. This was inferred from the relatively equal speed performance of buses and passenger

cars (note previously mentioned analysis and Figs. 2 and 3). The conclusion, therefore, was that buses do not constitute a deterrent to traffic flow on downgrades. Furthermore, since the passenger car equivalencies of loaded trucks on steep downgrades may assume values of 10 or more, previously suggested lower values may be questioned.

TABLE 5.—Passenger Car Equivalents of Loaded Trucks, Empty Trucks, and Buses

Downgrade site (1)	Average slope, in percent (2)	Length, in meters (3)	Calculated Passenger Car Equivalents: Loaded trucks (4)	Empty trucks (5)	Buses (6)
Atlit (level stretch)	0	—	2.07	2.07	1.0
Carmiel 1	9.0	500	10.6	2.4	3.3
Carmiel 2	4.5	530	3.9	2.0	0.9
Bat-Shlomo	4.8	500	5.7	3.2	—
Lavie	3.5	850	5.7	3.2	2.3
Elifelet	6.8	950	6.8	3.2	1.7
Migdal Ha'emek	7.7	1,150	7.0	2.0	1.2

Conclusions and Recommendations

The purpose of this research was to evaluate actual speed characteristics of heavy vehicles on downgrades. To that end, speed field observations were made on several rural downgrades. It was found that loaded trucks considerably reduce their speed at the beginning of a downgrade. The amount of reduction, termed the approach-speed gradient, was shown to be related to the length and slope of the downgrade, the second variable contributing exponentially to the increase in gradient. Empty trucks and buses were found to increase their speeds at the downgrades studied.

Based on a previously developed model, downgrade passenger car equivalents were calculated for loaded and empty trucks and for buses. The values for loaded trucks were found to be the highest; all downgrade values were higher than values for level operation. The values presented in this study and its method of calculation may be considered for further studies which will lead to development of a comprehensive set of passenger car equivalents on downgrades.

Further research is suggested: (1) On other rural downgrades with an additional variety of slopes and greater lengths; (2) on the flow characteristics of heavy vehicles on multi-lane, downgraded highways; and (3) on methods for improved designs of downgrades, based on practical application of such research.

Appendix I.—References

1. "A Policy on the Geometric Design of Rural Highways," American Association of State Highway Officials (AASHO), Washington, D.C., 1965.
2. Craus, J., Polus, A., and Grinberg, I., "A Revised Method for the Determination

of Passenger Car Equivalencies," *Transportation Research,* Vol. 14A, No. 4, Aug., 1980, pp. 241–246.
3. *Highway Capacity Manual,* Special Report 87, Highway Research Board, 1965.
4. Kmenta, J., *Elements of Econometrics,* The MacMillan Company, New York, N.Y., 1971.
5. Kobett, D. R., et al., "Traffic Simulation for the Design of Uniform Service Roads in Mountainous Terrain, Volume II: Description and Validation of the Simulation," Midwest Research Institute, Kansas City, Mo., 1970.
6. *Policy for Geometric Design of Rural Roads,* National Association of Australian State Road Authorities (NASRA), 4th ed., 1970.
7. Webb, G. M., "Downhill Truck Speeds," Traffic Bulletin No. 1, California Division of Highways, July, 1961.

Appendix II.—Notation

The following symbols are used in this paper:

A, B = calibration constants;
e = base of natural logarithm;
F = statistics which represent the F distribution;
f_1, f_2 = functions;
$\overline{GV}$ = mean approach-speed gradient for certain class of vehicles;
GV_k = approach-speed gradient for vehicle k;
i = average slope of downgrade;
k = vehicle index;
L = length of downgrade;
n = number of vehicles of given class at one site;
R = coefficient of determination;
$\bar{V}$ = average spot speed;
V_{k1} = spot speed for vehicle k at section 1 (approach speed);
V_{k2} = spot speed for vehicle k at section 2; and
σ = standard deviation of spot speeds.

TRANSPORTATION
ENGINEERING JOURNAL

Time-Dependent Route Assignment of Peak Traffic

By Attahiru Sule Alfa[1]

(Reviewed by the Urban Transportation Division)

Introduction

In most traffic assignment problems it is usually assumed that the rate at which vehicles leave one origin (O) for a destination (D) is given and the task is to assign traffic to different alternative routes between the O|D pair. There are, however, situations in which only the times at which they want to arrive at the D are known; let us call this time the destination target time (DTT). For the travelers, both their departure times from O and route selection depend on the DTT. A typical example of this is of commuters who wish to arrive at work at particular times. The aim of this paper is to develop a method for predicting route assignment when only the DTT and the total demand, between the O|D pair, are known and the departure rate is not known. The method developed is also capable of being used to predict their departure rate. This problem was first considered by Hurdle (3).

Hurdle (3) used a deterministic queueing model to find the traffic assignment on the assumption that each traveler selects his route and departure time from O in such a way so as to minimize his own time losses. Although Hurdle considered the importance of different drivers' willingness to gamble on being late or the difference in the relative values they give to time spent in queue as opposed to early arrival at their destination, he did not incorporate them in his model. These aspects will be included in the model being considered in this paper.

In this paper it is assumed that a traveler does not want to arrive at his destination late or too early and also he does not want to spend too much time in the system. It is then assumed that the traveler therefore selects both his route and departure time from O so as to minimize his total cost, on the

[1] Lect., Dept. of Civ. Engrg., Ahmadu Bello Univ., Samaru-Zaria, Nigeria.

Note.—Discussion open until August 1, 1981. To extend the closing date one month, a written request must be filed with the Manager of Technical and Professional Publications, ASCE. Manuscript was submitted for review for possible publication on June 17, 1980. This paper is part of the Transportation Engineering Journal of ASCE, Proceedings of the American Society of Civil Engineers, ©ASCE, Vol. 107, No. TE2, March, 1981. ISSN 0569-7891/81/0002-0153/$01.00.

assumption that he attaches some perceived costs to early and late arrivals and to delays in the system.

This paper considers only the case when all the travelers have the same DTT and O/D pair—although the result can be modified to incorporate the case when the travelers have different DTTs or O/D, or both. Therefore, it will be assumed that the travelers behave identically and independently; by considering their behavior and the whole system as stochastic then the variability of a traveler's behavior at different times and of different travelers is thus considered.

In this paper we shall use a stochastic approach to determine the combined temporal and route assignment of traffic during the morning peak period. We shall first develop a theoretical bivariate stochastic model for this problem, and later approximate the results by a "semistochastic" model for practical purposes. This approximation shall combine a stochastic model for the temporal distribution with a deterministic model for route assignment.

Theoretical Model

The major assumption in this model is that a commuter selects both his route and departure time from home in order to arrive at his destination on time by leaving home as late as possible. This model is an extension of the stochastic model for the temporal distribution of traffic demand (SMTD) by the writer and Minh (2), with two decision variables instead of one. We adopt the discrete time approach as in Ref. 2. If we consider N time epochs only and there are R routes in the network, then a commuter must now choose from $N \times R$ variables where previously in the SMTD he chose from N variables.

We shall assume that each of the R routes is clearly defined and that each route r consists only of a major link and a major bottleneck, which has the capacity (saturation flow) of $1/S^r$ vehicles per unit time (compare with Ref. 2). If t_r^n is the travel time on the route r if a commuter started his journey at epoch n, then

$$t_r^n = T_0^r + {}_r w^n \quad \text{. . . . .} \quad (1)$$

in which T^r = the minimum travel time on this route r; and ${}_r w^n$ = the delay encountered by the commuter who started his journey on this route at epoch n. Because the minimum travel time may be different for different routes we shall not ignore the value T_0^r as done in Ref. 2. We shall therefore consider departure time from home instead of arrival time at the bottleneck, in this paper.

We shall also assume that there are $I + 1$ commuters using this system. Except for the number of routes that has increased from one to R, every other assumption that applies to the SMTD, applies to the assignment model we are about to develop.

Let us define $\nu(d)$ as the route selected by a commuter on the dth day and let $\tau(d)$ be the epoch at which he departed from home on that day. Consider the bivariate process $[\tau(d), \nu(d); d \geq 1]$; our aim is to find a commuter's decision for each day d, i.e.:

$$\pi_n^r(d) \stackrel{\Delta}{=} \Pr\,[\tau(d) = n, \nu(d) = r] \quad \text{. . . . .} \quad (2)$$

in which $0 \leq \pi_n^r(d) \leq 1$; $\Sigma_{r=1}^{R} \Sigma_{n=1}^{N} \pi_n^r(d) = 1.0$; and $\pi_n^r(d)$ = the joint probability of a commuter's departure time from home and route selection for the dth day. Our interest is in the limiting distribution of $\pi_n^r(d)$, as $d \to \infty$, and also the marginal distribution $\alpha_r \triangleq \Sigma_{n=1}^{N} \pi_n^r(\infty)$, which is the probability of a commuter choosing route r. First, let us extend the theory of the SMTD to obtain $\pi_n^r(d)$.

Since it has been assumed that all the $I + 1$ commuters behave in a similar and independent manner, we shall only study the behavior of the typical commuter whom we name A, and who is one of these $I + 1$ commuters (compare to Ref. 2). Suppose A wishes to arrive at his destination at epoch T. Given that he departed from home at epoch $\tau = n$, used route $\nu = r$ and got delayed ${}_rw^n = i$ units of time, then he would arrive at his destination at epoch $n + T_0^r + i + S^r$, in which case one of the following events would occur: (1) $n + T_0^r + i + S^r < \mathrm{T}$; and he is early by an amount $\mathrm{T} - n - T_0^r - i - S^r$ to which he would attach a cost $C_e^r(n,i)$; (2) $n + T_0^r + i + S^r > \mathrm{T}$, and he is late by an amount $n + T_0^r + i + S^r - \mathrm{T}$ to which he would attach a cost $C_l^r(n,i)$; and (3) $n + T_0^r + i + S^r = \mathrm{T}$, which is the desirable event. He also attaches a cost $C_w^r(i)$ for spending $T_0^r + i + S^r$ units of time to travel.

The total cost incurred by this commuter A, $C^r(n,i)$, under the foregoing conditions, is thus given as;

$$C^r(n,i) \triangleq C_w^r(i) + C_e^r(n,i), \quad \text{if} \quad n + T_0^r + i + S^r < \mathrm{T};$$

$$C^r(n,i) \triangleq C_w^r(i) + 0, \quad \text{if} \quad n + T_0^r + i + S^r = \mathrm{T};$$

$$C^r(n,i) \triangleq C_w^r(i) + C_l^r(n,i), \quad \text{if} \quad n + T_0^r + i + S^r > \mathrm{T} \quad \ldots\ldots\ldots\ldots \quad (3)$$

Most assignment techniques usually consider only the cost due to travel time, i.e., $C_w^r(i)$, because the desired time of arrival at a destination is not considered to be a major factor for the kind of traffic considered. Hurdle (3) in the "work-to-home" model also considered just $C_w^r(i)$, however for the "home-to-work" model the inclusion of $C_l^r(n,i)$ was implied although he more or less assumed it to be infinity, for all $n + i + T_0^r + S^r > \mathrm{T}$. This assumption for the home-to-work model is not completely true because some commuters still arrive at work late. Hurdle, however, went further to consider the importance of considering the tradeoffs between arrival times at work (late or early) and delays (excess travel time) in the system.

Given that the delay encountered by A on the dth day, ${}_rw^n(d) = i$, if he used route r and departed from home at epoch n, let ${}_rW_i^n(d) \triangleq \Pr\,[{}_rw^n(d) = i]$.

Suppose on the dth day A departed from home at epoch n and used route r. If he got delayed i units of time then he would incur a total cost $C^r(n,i)$. If for the $(d + 1)$th day he wishes to reduce his total cost, and therefore changes his departure time from home to epoch m and uses route s. If for this $(d + 1)$th day he gets delayed j units of time, then by the same sort of arguments used in Ref. 2, the reduction in cost associated with this change is given by $[C^r(n,i) - C^s(m,j)]^+$, in which $(\delta)^+ = \max(0,\delta)$. Let $q_{n,m}^{r,s}(d)$ be the expected reduction in cost associated with A changing his route from route r and departure time from home from epoch n on the dth day to route s and epoch m on the $(d + 1)$th day, respectively, then, $q_{n,m}^{r,s}(d)$ is given by:

$$q^{r,s}_{n,m}(d) = \sum_{i=0}^{I\times S^r} \sum_{j=0}^{I\times S^s} [C^r(n,i) - C^s(m,j)]^+ {}_r W^n_i(d) {}_s W^m_j(d+1) \quad \ldots\ldots\ldots (4)$$

A's choice of departure time from home and route for the $(d+1)$th day, $[\tau(d+1), \nu(d+1)]$, depends on the values of $q^{r,s}_{n,m}(d)$, which he estimates by $\bar{q}^{r,s}_{n,m}(d)$ in which, based on the same arguments as for the SMTD in Ref. 2:

$$\bar{q}^{r,s}_{n,m}(d) = \sum_{i=0}^{I\times S^r} \sum_{j=0}^{I\times S^s} [C^r(n,i) - C^s(m,j)]^+ {}_r W^n_i(d) {}_s W^m_j(d) \quad \ldots\ldots\ldots (5)$$

Let $t^{r,s}_{n,m}(d) \triangleq \Pr[\tau(d+1) = m, \nu(d+1) = s \mid \tau(d) = n, \nu(d) = r]$. Letting $b^r_n(d) \triangleq \Sigma^R_{u=1} \Sigma^N_{v=1} \bar{q}^{r,u}_{n,v}(d)$, and using arguments similar to those for the SMTD, we obtain the transition probability $t^{r,s}_{n,m}(d)$ as

$$t^{r,s}_{n,m}(d) = \frac{\bar{q}^{r,s}_{n,m}(d)}{b^r_n(d)}, \quad \text{if} \quad b^r_n(d) > 0,$$

$$t^{r,s}_{n,m}(d) = 1, \quad \text{for} \quad r = s \quad \text{and} \quad n = m, \quad \text{if} \quad b^r_n(d) = 0;$$

$$t^{r,s}_{n,m}(d) = 0, \quad \text{otherwise}, \quad \text{if} \quad b^r_n(d) = 0 \quad \ldots\ldots\ldots (6)$$

Since all the other I commuters behave in the same way as A does we shall now drop the name label and consider the system either as seen by an arbitrary commuter or as seen by all the $I + 1$ commuters.

Let us define an $(N \times R) \times (N \times R)$ transition matrix $\mathbf{T}^R(d)$ such that $[\mathbf{T}^R(d)]_{r\times N+n, s\times N+m} \triangleq t^{r+1,s+1}_{n,m}(d)$; $0 \le (r,s) \le R - 1$; $1 \le (n,m) \le N$; and also define an $(N \times R)$ vector $\mathbf{\Pi}^R(d)$, such that $\mathbf{\Pi}^R(d) \triangleq [\pi^1_1(d) \ldots \pi^1_N(d)\, \pi^2_1(d) \ldots \pi^2_N(d) \ldots \pi^R_1(d) \ldots \pi^R_N(d)]$, then a commuter's decision as to which epoch to depart from home and on which route to travel on the $(d+1)$th day can be set up as a two-dimensional Markov Chain with a time-inhomogeneous transition matrix that incorporates a feedback mechanism as

$$\mathbf{\Pi}^R(d+1) = \mathbf{\Pi}^R(d) \times \mathbf{T}^R(d) \quad \ldots\ldots\ldots (7)$$

As in the SMTD, we assume that there exists a limiting distribution $\mathbf{\Pi}^R = \lim_{d\to\infty} \mathbf{\Pi}^R(d)$ which coincides with $\mathbf{T}^R = \lim_{d\to\infty} \mathbf{T}^R(d)$. Our interest is to find $\mathbf{\Pi}^R$ which gives us the information we require about the combined temporal distribution and route assignment of traffic over the network.

This model is an extension of the SMTD, and their algorithms are therefore similar. However, due to the large dimensionality of the present model its computation time and storage requirements are astronomically large—on the order of R^2 of those required for just the temporal distribution of traffic demand on one route, considering only one iteration in each case. We shall therefore use some heuristic judgment, combined with some known properties of the assigned traffic, to approximate the foregoing algorithm for practical purposes.

Approximate Procedure for Assignment of Peak Period Traffic

Let us assume that the number of commuters, $I + 1$, involved in this system is quite large. Let us assume, for the time being, that we know the values

of α_r ($r = 1, 2, \ldots, R$), in which α_r = the probability that a commuter uses route r. If, of the total population $I + 1$ using the system, I_r commuters among them use route r, then the law of large numbers ensures that

$$\alpha_r \simeq \frac{I_r}{I+1} \qquad (8)$$

For practical purposes let us assume that I_r is an integer. By the preceding arguments, if there exists a steady-state solution for Eq. 7, then I_r is fixed and constant. We can then assume that I_r is the total population of commuters who have decided to use route r after having tried all the other routes. It therefore follows that these I_r commuters will always use route r and select their departure times from home as in the SMTD. Let us call the departure process obtained via the SMTD the suboptimal departure process (SODP). If we know the value of I_r then we can use the SMTD to obtain this departure process for route r; let this SODP for this route r used by the I_r commuters be $\tilde{\Pi}_r$, with $\tilde{\pi}^r_n \stackrel{\Delta}{=} [\tilde{\Pi}_r]_n$, $0 \le \tilde{\pi}^r_n \le 1, \forall n$, r and $\Sigma^N_{n=1} \tilde{\pi}^r_n = 1$. If we partition Π^R in Eq. 7 and let $\Pi^R = [\Pi_1|\Pi_2|\ldots|\Pi_R]$, in which $\Pi_r \stackrel{\Delta}{=} [\pi^r_1 \pi^r_2 \ldots \pi^r_N]$, then it is reasonable to assume that Π_r can be approximated by

$$\Pi_r \simeq \alpha_r \times \tilde{\Pi}_r \qquad (9)$$

keeping in mind that $0 \le \alpha_r \simeq I_r/(I + 1) \simeq \Sigma^N_{n=1} \pi^r_n \le 1, \forall r$. The problem then is how to obtain I_r.

Let us define $EC_r(I_r)$ as the average expected total cost incurred by each of the I_r commuters using route r under the SODP $\tilde{\Pi}_r$ then

$$EC_r(I_r) = \sum_{n=1}^{N} \tilde{\pi}^r_n \sum_{i=0}^{I_r \times S^r} C^r(n, i)_r \tilde{W}^n_i \qquad (10)$$

in which $_r\tilde{W}^n_i$ = the delay distribution associated with $\tilde{\Pi}_r$.

Let us relabel the routes in their order of "attractiveness," such that:

$$EC_1(k) < EC_2(k) < \ldots < EC_{R-1}(k) < EC_R(k), \forall k \ge 1, \qquad (11)$$

and ignore the trivial case of $EC_r(k) = EC_s(k), \forall r \ne s$, for the time being. This labeling is for convenience and not essential.

As mentioned earlier, a commuter's ultimate aim is to select a route and a departure process that he thinks will minimize his total cost. It is therefore evident from Eq. 11 that if there is only one commuter using the system then he will definitely use route 1. In which case he will incur an expected total cost of EC_1 (1), with his departure process, in this case, being deterministic.

Suppose a second commuter is introduced into the system. Let us assume that this second commuter behaves in the same manner but independent of the first commuter. This second commuter will also probably think of using route 1 because he anticipates incurring the least cost if he does. If both commuters use route 1, then they will have to select a departure process for which they will incur an expected total cost of EC_1(2) each. An important question either one or both of these commuters would probably ask himself is whether it is possible for him to use route 2 and find a departure process for which the expected total cost he would incur would be less than EC_1(2). The two commuters

would therefore be left with two options viz, either: (1) They both use route 1, if $EC_1(2) < EC_2(1)$, or; (2) one uses route 1 and the other uses route 2, if $EC_1(2) > EC_2(1)$. Thus even by this approximation method, traffic assignment is still a bivariate decision-making process, with both the route and departure time as the decision variables. Note that when $EC_1(2) = EC_2(1)$, both the foregoing solutions are equally reasonable and likely. When such situations arise in real life, the other intangible system variables would usually dictate a commuters' preference.

This strategy can now be extended to the case involving all the $I + 1$ number of commuters and all the R routes. Using the same arguments as before it can be shown that all these $I + 1$ commuters will select their routes in such a manner that:

$$EC_r(I_r) \leq EC_{r+1}(I_{r+1}); \quad r = 1, 2, \ldots, R - 1 \quad (12)$$

and $$EC_r(I_r + 1) > EC_{r+1}(I_{r+1}) \quad (13)$$

keeping in mind that

$$I + 1 = \sum_{r=1}^{R} I_r \quad (14)$$

Conditions in Eqs. 12 and 13 merely state that the commuters assign themselves to all the routes such that the average expected total cost to the commuters using route r is less than or equal to the average expected total cost to those using route $r + 1$, provided Eq. 11 holds. However, although the average expected total cost to those using route $r + 1$ could be reduced if one commuter switches from this route to route r, no commuter will make this switch because if he did he would incur more cost than he would have had he remained on route $r + 1$; on the other hand if a commuter switched from route $r + 1$ to route r under such conditions, then another commuter would have to switch to route $r + 1$ from route r due to the increase in cost that he notices. This is conceptually similar to the user equilibrium assignment criterion, with travel time replaced by the average expected total cost. Another way of looking at this is that since route r is more "attractive" than route $r + 1$, at least under equal flow conditions, a commuter will always prefer to use route r. However, he will switch to route $r + 1$ if, and only if, the cost he will incur when using route $r + 1$ is going to be less than the cost he would have incurred had he remained on route r.

Basically the proposal of what a commuter goes through in selecting his route, even according to this approximation technique, is not independent of the time he departs from home, which dictates his arrival time at his destination. Initially, he arbitrarily selects a route and a departure time. He then changes his departure times on many occasions on this route until he achieves the SODP which he thinks "minimizes" his expected total cost. However, if he is still not satisfied with this minimum cost he incurs on this route, he will then try another route. On this new route he will search for another SODP to which there will be another expected total cost attached. After trying many routes and seeking the "best" departure process on each of these routes, a commuter finally settles for that route on which he incurs the "least" expected total cost. These arguments led to Eqs. 12 and 13. This procedure uses a stochastic procedure for selecting departure time from home and a deterministic approach for selecting route.

For the assignment procedure, we are looking for a unique set of integer values of $I_r, \forall r$, that satisfy Eqs. 12, 13, and 14 and the associated SODP.

Computational Procedure for Assignment

The computational procedure for this assignment is fairly similar to that used for capacity restraint. The first step in this procedure is to consider one route r separately and load some known number, I_r of commuters on to this route. Using the SMTD, obtain the "best" temporal distribution of these I_r commuters' departure times from home while using this route (i.e., the SODP), which is given by $\tilde{\Pi}_r$; and also obtain the average expected total cost $EC_r(I_r)$, associated with this distribution. Then increase I_r and repeat the foregoing until $I_r \to I + 1$, depending on how competitive all the routes are. For each value of I_r there is an associated vector $\tilde{\Pi}_r$ and a value $EC_r(I_r)$ which have both now been computed. This process is further carried out for all the other $R - 1$ routes in the system. An $(I + 1) \times R$ table of I_r versus $EC_r(I_r)$, $1 \le I_r \le I + 1, \forall r$, can thus be set up. In some circumstances, a graphical representation of these might be more convenient for usage.

The next step is to consider the first route ($r = 1$) and assign all the $I + 1$ commuters to this route. Then test and see if $EC_1(I + 1) \le EC_2(1)$. If this is true, then all the $I + 1$ commuters would use route 1 and the other $R - 1$ routes will not be used at all. However, if $EC_1(I + 1) > EC_2(1)$, then the $I + 1$ commuters are assigned to both routes 1 and 2. This is done using the equality condition that $EC_1(I_1) = EC_2(I_2)$, if a graphical method is employed, keeping in mind that $I + 1 = I_1 + I_2$. Having obtained the values of I_1 and I_2, which might not be integer values, they are rounded off to satisfy Eqs. 12, 13, and 14. However, if the direct figures from the table are used, then I_1 and I_2 are chosen by a trial and error procedure until Eqs. 12, 13, and 14 are satisfied. A graphical method is probably simpler and faster, particularly when the values of $I + 1$ and R are large. It was discovered, however, that when all the three cost parameters are linear and nonzero, in most cases considered the relationship between I_r and $EC_r(I_r)$ is linear and an analytical method can be used to solve the assignment problems. Anyway, having assigned the traffic to routes 1 and 2, we then test whether $EC_2(I_2) \le EC_3(1)$. If this is true then the assigned traffic remain as before, otherwise all the traffic are assigned to the three routes as done previously such that $EC_1(I'_1) = EC_2(I'_2) = EC_3(I'_3)$, in which I'_r = the number of commuters assigned to route r at this stage, keeping in mind that $I + 1 = I'_1 + I'_2 + I'_3$. This process continues until all the $I + 1$ commuters have been assigned to all the R_u routes being used, in which $R_u \le R$. The final value of I_r assigned to a route r denotes the number of commuters using that route, and the corresponding $\tilde{\Pi}_r$ is the distribution vector of these commuters' departure times from home. Before we provide an illustrative example, let us briefly consider an important aspect of the cost parameters.

Cost Parameters

It was explained in Ref. 1 that it is a reasonable approximation to assume that the cost parameters are all linear for most commuters. There we let:

$$C^r(n,i) = C^r_w \times (i) + C^r_e \times (\mathrm{T} - n - T^r_0 - i - S^r), \quad \text{if} \quad n + T^r_0 + i + S^r \le \mathrm{T};$$

$$C^r(n,i) = C^r_w \times (i) + C^r_l \times (n + T^r_0 + i + S^r - \mathrm{T}),$$

$$\text{if} \quad n + T^r_0 + i + S^r \ge \mathrm{T} \quad \ldots \ldots \ldots \ldots \ldots \ldots \quad (15)$$

If we consider only the cost for delay and assume that $C^r_e = C^r_l = 0.0$, it is well known that even if C^r_w is a constant the expected total cost $EC_r(I_r)$ is convex and increases monotonically in a nonlinear manner as the total population, I_r, increases. This is further shown in Fig. 1 where $C^r_w = 1.0$ and $C^r_e = C^r_l = 0.0$. However, when $C^r_e > 0$, $C^r_l > 0$, and $C^r_w > 0$ and they are all constant we find that the cost for earliness or lateness, as the case may be, "normalize" the cost due to delay to produce a linear relationship between $EC_r(I_r)$ and I_r. The normalizing in this context implies that for those travelers who incur high delay costs we would expect them to do so only if they expect

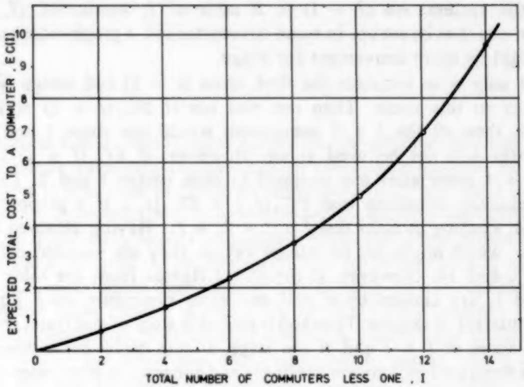


FIG. 1.—Relationship between $EC(I)$ and I, when Cost for Delay Is Linear and Greater than Zero and Other Two Costs Are Zero

to incur low cost for either lateness or delay and vice versa. The reason being that these costs only represent tradeoffs and act as balancing factors. This postulate is further supported by some simulation results where nonzero linear cost functions are assumed with different rates and the number of commuters varied. The results are presented in Fig. 2. For curve 1 the costs for earliness, lateness and delay were assigned the rates of 1.0, 2.0, and 0.5, respectively, and for curve 2 these costs were assigned the rates of 1.0, 1.0, and 6.5, respectively. The units of these costs can be assumed to be, say, dollars per unit time. In both cases the relationship between $EC_r(I_r)$ and I_r is linear, as expected.

We therefore suggest that for the aforementioned kind of problem $EC_r(I_r)$ can be approximately represented for each route r in the form:

$$EC_r(I_r) = \beta^0_r + \beta^1_r \times I_r \quad \ldots \ldots \ldots \ldots \ldots \ldots \quad (16)$$

in which the two coefficients β^0_1 and β^1_r can be obtained by first obtaining

similar curves as in Fig. 2 for each route, then running a simple linear regression analysis to estimate the best fitting values of these coefficients. The term β_r^0 is a function of the minimum travel time on route r and the cost associated with this value, i.e., C_w^r. The term, β_r^1, is a function of the commuter's arrival time at the destination, which depends on the capacity of the route r he travels on, thus it depends on all the three cost parameters. Let us use some hypothetical examples to illustrate this algorithm.

Illustrative Examples

For illustrative purpose let us consider a system consisting of two routes ($R = 2$) between an origin O and a destination D. Suppose there are 1,000 commuters wanting to go from O to D and that they all want to arrive at

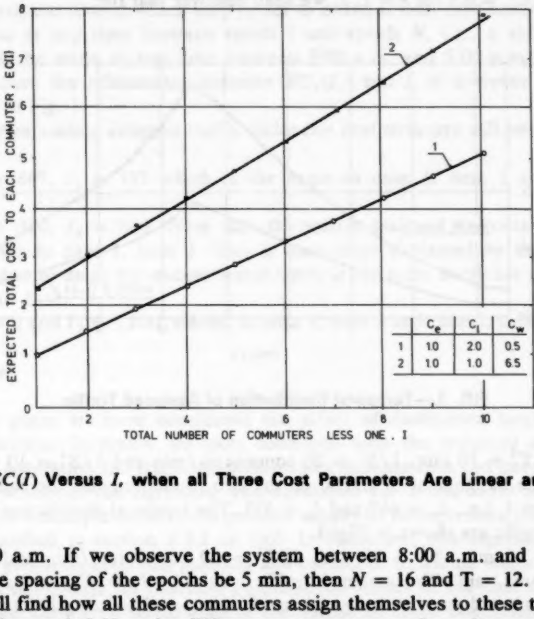


FIG. 2.—$EC(I)$ Versus I, when all Three Cost Parameters Are Linear and Greater than Zero

D at 9:00 a.m. If we observe the system between 8:00 a.m. and 9:20 a.m. and let the spacing of the epochs be 5 min, then $N = 16$ and $T = 12$.

We shall find how all these commuters assign themselves to these two routes over the time period N, under different cost structure and varying route factors.

Case 1.—Suppose that all these commuters attach linear costs to earliness, lateness and delay with the rates 1.0, 1.0, and 6.5 respectively; and assume that these rates are the same for both routes. If under this cost structure:

1. The minimum travel time on both routes are the same and equal to 10 min, ($T_0^1 = T_0^2 = 10$ min); the capacity of route 1 is 20 commuters/min and the capacity of route 2 is 10 commuters/min, then the number of traffic assigned to each route will be given by $I_1 = 667$ and $I_2 = 333$.

2. Everything remains the same as in item 1 except that the minimum travel time on route 1 is increased to 15 min. The assigned traffic now becomes $I_1 = 305$ and $I_2 = 695$. This considerable change in the assigned traffic is brought about by the fact that the most sensitive cost parameter is the cost for delay (time in the system), C_w^r, which is quite dominant. Therefore, increasing T_0^1 from 10 min–15 min makes the average total cost to a commuter, using route 1, go up by a reasonable amount, thereby making this route very much less attractive.

3. We further increase the capacity of route 2 to 20 commuters/min and let everything else remain the same as in item 2, then route 2 gets so much more attractive than route 1 that all the 1,000 commuters will switch to using only route 2, i.e. $I_1 = 0$ and $I_2 = 1{,}000$.

Case 2.—Suppose that we change the rates of the costs to $C_e^r = 1.0$, $C_l^r = 2.0$ and $C_w^r = 0.5$ for $r = 1, 2$, we shall discover that for:

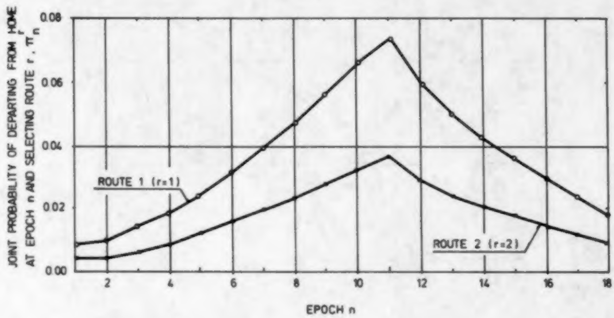


FIG. 3.—Temporal Distribution of Assigned Traffic

1. $T_0^1 = T_0^2 = 10$ min, $1/S^1 = 20$ commuters/min and $1/S^2 = 10$ commuters/min, the assignment of traffic on the two routes remains the same as for case 1, item 1, i.e., $I_1 = 667$ and $I_2 = 333$. The temporal distribution of these assigned traffic are shown in Fig. 3.

2. $T_0^1 = 15$ min, $T_0^2 = 10$ min, $1/S^1 = 20$ commuters/min and $1/S^2 = 10$ commuters/min. The assigned traffic would be $I_1 = 630$ and $I_2 = 370$. The assignment does not change as much, compared to case 1, item 2. This is because in this present example the cost to time in the system, C_w^r is not dominant in the cost structure.

3. $T_0^1 = 15$ min, $T_0^2 = 10$ min, $1/S^1 = 1/S^2 = 20$ commuters/min, all the commuters will switch to using route 2 as in case 1, item 3.

In both case 1, item 3 and case 2, item 3, all the 1,000 commuters were able to choose their departure times from home and use route 2 and still incur less total cost than they would have had they used route 1 and even incurred the minimum possible cost. However, if the total number of the commuters

is increased to 2,000 we shall discover that some of them will than have to use route 1. In fact, if this happens then under case 1: $I_1 = 458$ and $I_2 = 1{,}542$ and under case 2: $I_1 = 945$ and $I_2 = 1{,}055$.

In this assignment model we have shown that if most of the traffic system users during the peak period are commuters then it is very important to include, in the factors that affect their route choice, the desire for them to arrive at work at a particular time which does affect the temporal distribution of their demand. Secondly, we have shown that the temporal distribution of work trip generation and commuters' route choice during the peak period are interdependent.

In the next example we shall show the difference in our results if we ignore both the cost for late and early arrivals.

Case 3.—Suppose that we change the rates of the costs to $C'_e = C'_l = 0.0$ and $C'_w = 1.0$. This is equivalent to the situation when the commuters do not have a particular time at which they prefer to arrive at their destination provided they arrive at any time between epoch 1 and epoch N, i.e., a shopping trip (that could be made at any time between 9:00 a.m. and 5:00 p.m.). For this cost structure the relationship between $EC_r(I_r)$ and I_r is however not linear as shown in Fig. 1.

The corresponding assigned traffic under this cost structure will be as follows:

1. $I_1 = 667$, $I_2 = 333$ which is the same as case 1, item 1 and case 2, item 1.
2. $I_1 = 160$, $I_2 = 840$. Note that the traffic assigned to route 2 is even greater than in case 1, item 2. This is once more explained by the fact that the cost due to delay (or excess travel time) is the most dominant or the only cost in this case.
3. $I_1 = 0$ and $I_2 = 1{,}000$, similar to case 1, item 3 and case 2, item 3.

Conclusion

In this paper we have considered the effect of destination target time on route selection. In reality we have combined both the temporal distribution of traffic generation and traffic assignment as a simultaneous and interdependent process. Although the algorithm was developed for a one O/D pair, it can be used for a multiple destinations problem simply by incorporating the modifications described in section 3.3.2 of Ref. 1. Therefore, this algorithm is quite suited to problems involving planning and control of traffic on major routes.

To use this model for practical purposes the values of the cost coefficients have to be known. Whereas finding the values of these coefficients in itself is quite involved, they can be obtained as calibration factors for a known traffic situation before such traffic situation can be analyzed. These factors have been obtained in Ref. 1 for both the southbound and northbound traffic using the Sydney Harbor Bridge during the morning peak period.

Appendix I.—References

1. Alfa, A. S., "A Model for the Temporal Distribution of Peak Traffic Demand—Development, Calibration and Application to Route Assignment," thesis presented to the

University of New South Wales, at Kensington, N.S.W. Australia, in 1979, in partial fulfillment of the requirements for the degree of Doctor of Philosophy.
2. Alfa, A. S., and Minh, D. L., "A Stochastic Model for the Temporal Distribution of Traffic Demand—the Peak Hour Problem," *Transportation Science*, Vol. 13, No. 4, Nov., 1979, pp. 315–324.
3. Hurdle, V. F., "The Effect of Queueing on Traffic Assignment in a Simple Road Network," *Proceedings of the 6th International Symposium on Transportation and Traffic Theory*, D. J. Buckley, ed., pp. 519–540.

Appendix II.—Notation

The following symbols are used in this paper:

$b^r_n(d)$ = $\Sigma^N_{m=1} \Sigma^R_{s=1} \bar{q}^{r,s}_{n,m}(d)$;
$C^r(n,i)$ = commuter's total cost given that he departed from home at epoch n, used route r and got delayed i units of time;
$[C^r(n,i) - C^s(m,j)]^+$ = max $[0, C^r(n,i) - C^s(m,j)]$;
C^r_e = cost coefficient for earliness on route r;
$C^r_e(n,i)$ = commuter's cost for being early at his destination by amount $T - n - i - S^r - T^r_0$ given that he departed from home at epoch n, used route r, and got delayed i units;
C^r_l = cost coefficient for lateness on route r;
$C^r_l(n,i)$ = commuter's cost for being late at his destination by amount $n + i + S^r + T^r_0 - T$, given that he departed from home at epoch n, used route r, and got delayed i units;
C^r_w = cost coefficient for delay on route r;
$C^r_w(i)$ = commuter's cost for being delayed i units and spending $T^r_0 + S^r + i$ units of time to travel, given that he used route r;
$EC_r(I_r)$ = expected total cost to commuter, given that he uses route r with $I_r - 1$ other commuters;
$I + 1$ = total number of commuters using system;
I_r = total number of commuters using route r;
I'_r = number of commuters assigned to route r during one iteration;
N = total number of time epochs being considered;
$q^{r,s}_{n,m}(d)$ = total expected reduction in cost incurred by commuter if he changed his departure time from home and route from epoch n and route r on dth day to epoch m and route s on $(d + 1)$th day;
$\bar{q}^{r,s}_{n,m}$ = estimated volume of $q^{r,s}_{n,m}(d)$;
R = number of routes that exist between O|D pair;
R_u = number of routes being used by commuters;
S^r = service time at bottleneck on route r;
T^r_0 = minimum travel time on route r;
t^n_r = traveler's travel time on route r, given that he started his journey at epoch n;

$t^{r,s}_{n,m}(d)$ = Pr $[\tau(d + 1) = m,\ \nu(d + 1) = s \mid \tau(d) = n,\ \nu(d) = r]$;

$_rW^n_i(d)$ = Pr $[_rw^n(d) = i]$;

$_r\tilde{W}^n_i(d)$ = delay distribution [i.e., $_rW^n_i(d)$] associated with $\tilde{\Pi}_r$;

$_rw^n(d)$ = delay to traveler on route r, given that he arrived at bottleneck at epoch n, on dth day;

α_r = Pr $[\nu(d) = r] = \Sigma^N_{n=1} \pi^r_n(d)$;

β^0_r, β^1_r = regression coefficients for linear relationship between total cost and number of commuters;

Π^R = $\lim_{d\to\infty} \Pi^R(d)$;

$\Pi^R(d)$ = $N \times R$ vector with its $(r \times N + n)$th element as $\pi^{r+1}_n(d)$;

$\Pi_r(d)$ = N vector with its nth element given as $\pi^r_n(d)$;

$\tilde{\Pi}_r, \tilde{\pi}^r_n$ = $\lim_{d\to\infty} (\Pi_r(d)/\alpha_r,\ \pi^r_n(d)/\alpha_r)$;

$\pi^r_n(d)$ = Pr $[\tau(d) = n,\ \nu(d) = r]$;

$\nu(d)$ = route selected by commuter on dth day;

$\mathbf{T}$ = destination target time (DTT);

$\mathbf{T}^R$ = $\lim_{d\to\infty} \mathbf{T}^R(d)$; and

$\mathbf{T}^R(d)$ = $(N \times R) \times (N \times R)$ matrix with element of $(r \times N + n)$th row and $(s \times N + n)$th column as $t^{r+1,s+1}_{n,m}(d)$.

[illegible]

TRANSPORTATION ENGINEERING JOURNAL

Dynamic Pavement Deflections

By Gary W. Sharpe,[1] Herbert F. Southgate,[2] and Robert C. Deen,[3] F. ASCE

(Reviewed by the Highway Division)

Introduction

In 1977, a methodology was developed to evaluate the behavior or performance of flexible pavements using dynamic (Road Rater) deflections (14). This method utilized deflections computed by elastic theory using the Chevron computer program (7,15) to simulate Road Rater deflections. This methodology involved a series of graphical interpolations and required considerable engineering judgment. Modifications presented in this paper refine and simplify the procedure considerably.

Nondestructive tests of pavements have been empirically correlated with field strength tests. There has been considerable use of elastic theory and dynamic testing to estimate layer moduli. Equipment used has included the Benkelman beam, the California traveling deflectometer, a falling-weight deflectometer, the Dynaflect, the Road Rater, and other vibratory testers (8).

Of the various devices available for the nondestructive measurement of pavement deflections, the simplest and most commonly used is the Benkelman beam. The Benkelman beam uses a lever to measure the rebound deflection as the tires of a test vehicle of known weight roll past the tip or "probe" of the beam (29). The beam has been successfully mechanized, making it possible to measure deflections continuously under a moving axleload. The California traveling deflectometer is one device utilizing the principle (8,29).

[1] Research Engr. Principal, Div. of Research, Bureau of Highways, Dept. of Transportation, 533 South Limestone, Lexington, Ky. 40508.

[2] Research Engr. Chf., Div. of Research, Bureau of Highways, Dept. of Transportation, 533 South Limestone, Lexington, Ky. 40508.

[3] Asst. Dir, Div. of Research, Bureau of Highways, Dept. of Transportation, 533 South Limestone, Lexington, Ky. 40508.

Note.—Discussion open until August 1, 1981. To extend the closing date one month, a written request must be filed with the Manager of Technical and Professional Publications, ASCE. Manuscript was submitted for review for possible publication on June 3, 1980. This paper is part of the Transportation Engineering Journal of ASCE, Proceedings of the American Society of Civil Engineers, ©ASCE, Vol. 107, No. TE2, March, 1981. ISSN 0569-7891/81/0002-0167/$01.00.

Benkelman beam deflections can be used to evaluate pavement performance; such analyses have been incorporated into overlay design procedures (12,29). Benkelman beam deflections have also been correlated with Road Rater deflections (16).

The second generation of measurement devices are the vibratory or impact testers, or both. Included in this group are the Dynaflect, the Road Rater, the WES 72.6-kN and 40.1-kN (16-kip and 9-kip) vibrators, and the Shell 18.1-kN (4-kip) vibrator (8). In general, vibratory testers induce a steady-state sinusoidal vibration in the pavement with a dynamic force generator. The magnitude of the dynamic force and the means by which the force is generated are the primary differences among vibratory testers. An impact device (the falling-weight deflectometer) measures the surface deflection resulting when a known weight is dropped a specified distance (8).

Deflection bowls have been used to study pavement response characteristics (3,5,6,10,11,25). Both empirical and theoretical analyses utilizing elastic theory have been considered. Factors considered in the evaluation of the deflection bowls, such as defined by dynaflect data, are maximum deflection, spreadability, surface curvature index, base curvature index, and the fifth sensor deflection. Spreadability, defined as the average deflection for all five Dynaflect sensors expressed as a percentage of the maximum deflection, is a measure of the ability of the pavement structure to distribute the load. The surface curvature index is the difference in deflections at the first and second sensors. It is a measure of the condition of the upper layers of the pavement structure. The base curvature index is the difference between the deflections at the fourth and fifth sensors and is meant to be an indication of the condition of the lower layers. Normally, the maximum deflection is an indicator of the condition of the bound layers while the deflection at the fifth sensor is an indicator of subgrade adequacy (3,5,6,10,11,25).

The falling-weight deflectometer also has been used to study pavement behavior. Procedures have been developed which incorporate falling-weight data into overlay designs for flexible pavements (4).

Kentucky research has used Road Rater deflections since 1972. Other organizations have used Road Rater deflections to evaluate the structural condition of flexible pavements (1,9). Procedures presented in this paper have evolved from several earlier studies (14,15,18,19,22,23) and represent those currently used in Kentucky. Research is continuing to further refine the evaluation process.

Simulation of Road Rater Deflections

Characteristics of Road Rater.—The testing head of the Kentucky Road Rater consists of a vibrating mass weighing 72.6 kg (160 lb) which impulses the pavement. The test head is lowered to the pavement until the hydraulic pressure of 4.82 MPa (700 psi) produces a static load of 7.428 kN (1,670 lb). The mass is vibrated using preselected frequencies of 10 Hz, 20 Hz, 25 Hz, and 40 Hz. The forced motion of the pavement is measured by velocity sensors normally located at 0 mm (0 ft), 305 mm (1 ft), 610 mm (2 ft), and 914 mm (3 ft) from the center of the test head.

Vibrating the mass at a frequency of 25 Hz and an amplitude of 1.524 mm (0.06 in.) results in a peak-to-peak dynamic force of 2.668 kN (600 lb). Once

the dynamic force is set for a given frequency and amplitude, the other preset frequencies will vary the amplitude of the vibrating mass such that the dynamic force remains fixed. The composite loading consists of a dynamic force of 2.668 kN (600 lb) amplitude oscillating about the static load of 7.428 kN (1,670 lb).

Superposition Principles.—The Road Rater loading is transmitted to the pavement by two "feet" symmetrically located on either side of a beam extending ahead and carrying the sensors. Superpositioning is applicable provided the deformations are small and do not substantially affect the action of external forces. A linear relationship between displacement and external force must exist or be assumed to exist (24). Applying superposition principles to the Road Rater, the deflection resulting from the load applied to one "foot" is added to the deflection due to the load applied by the other "foot." For the symmetrical conditions of the Road Rater, deflection calculations only need be made for one "foot" and the radii corresponding to each sensor location.

The dynamic loading (sine wave) of the Road Rater can be approximated by a square wave such that its amplitude is $1/\sqrt{2}$ times the amplitude of the sine wave. The maximum and minimum loadings of the square wave are 8.37 kN (1,882 lb) and 6.49 kN (1,458 lb). From symmetry, the loads on each "foot" of the test head are equal to 4.186 kN (941 lb) and 3.243 kN (729 lb). The dynamic deflection is defined by $D(\text{total}) = [D(4.19) - D(3.24)] \times 2$, in which $D(4.19)$ and $D(3.24)$ represent the deflections calculated by the Chevron computer program for the maximum and minimum loading conditions on one "foot."

Input Parameters for Chevron Computer Program.—In addition to load, required inputs to the Chevron program include a contact pressure corresponding to the: (1) Load; (2) number of layers; and (3) the thickness, Young's modulus, and Poisson's ratio for each layer. The contact pressure of the maximum and minimum loads are varied to maintain a constant area [102 mm × 178 mm (4 in. × 7 in.)] for each "foot." The following constants were used to calculate simulated Road Rater deflections (22):

1. Poisson's ratio—asphaltic concrete: $\mu = 0.40$; granular base: $\mu = 0.40$; and Subgrade: $\mu = 0.45$.
2. Load = 4.186 kN (941 lb): Contact Pressure = 0.231 MPa (33.5 psi).
3. Load = 3.243 kN (729 lb): Contact Pressure = 0.183 MPa (26.5 psi).

A matrix of asphaltic concrete (AC) thicknesses and moduli, dense-graded aggregate (DGA) thicknesses, and the constants indicated herein were used as input. Simulated deflections were calculated using elastic theory; these deflections were used to develop theoretical relationships presented in this paper (14,15).

Moduli of granular bases ($E2$) are a function of the moduli of the confining layers of asphaltic concrete ($E1$) and subgrade ($E3$). The modulus of the crushed stone layer is estimated from the relationship $E2 = F \times E3$, in which there is an inverse linear relationship between F and log $E3$. The ratio of the modulus of the base to the modulus of the subgrade is equal to 2.8 at a California bearing ratio (CBR) of seven and is equal to 1.0 when $E1$ equals $E3$, i.e., $E1 = E2 = E3$—the case of a Boussinesq semi-infinite half space (22). Subgrade moduli in pounds per square inch may be approximated by the product of the CBR and 1,500. This method of estimating base moduli appears adequate for normal design considerations up to a CBR of about 18 (2,14,27,28).

Reference Conditions.—The modulus of elasticity of asphaltic concrete varies as a function of frequency of loading and temperature. Conditions for the current Kentucky thickness design procedures and the method for conducting Benkelman beam tests correspond to a modulus of 3.31 GPa (480 ksi) at 0.5 Hz and a pavement temperature of 21.1° C (70° F) is 8.27 GPA (1,200 ksi) (19,20,21).

Because of the significant effects of temperature on modulus of elasticity of asphaltic concretes, a system was developed to adjust deflection measurements to a reference temperature and modulus. The adjustment scheme used ratios of deflections at reference conditions to deflections resulting from arrayed variables of layer thicknesses and moduli (14,15,16,18,20,21).

Evaluation of Pavement Structure

Foundation (subgrade) stiffness is a factor affecting the behavior of a pavement structure. Variations in subgrade support occur mainly as a result of variations in moisture contents and in soil types. A significant decrease in subgrade stiffness (or modulus of elasticity) will result in a loss of ability to support the pavement structure adequately and will lead to increased distress in the layers of the structure. Signs of distress include rutting, increased roughness, and cracking (29).

Estimates of subgrade strength are necessary to evaluate overall pavement conditions. A "design" condition exists when there is no loss of "effective" thickness in any of the layers. A knowledge of as-built (design) thicknesses of layers is necessary before an evaluation of the pavement structure can be made. Those thicknesses should be available from construction or maintenance records or both; cores also may be obtained to determine or verify layer thicknesses. Generally, conditions involve deterioration in the layers of the structure. This means that the individual layers are behaving similar to another combination of layer thicknesses composed of new-quality materials, i.e., the structure is behaving as an "effective" structure. In such a case it is necessary to estimate the "effective" thicknesses of the deteriorated structure.

Describing Shape of Deflection Bowl

The analysis of deflections involves the shape of the deflection bowl (3,5,6,10, 11,12,14,15,23,24,25,26). The Number 1 projected deflection, an empirical evaluation of Road Rater deflection data (14,15,16,23), is obtained by extrapolating a straight line through the deflection values of the Number 2 and Number 3 sensors when log deflection is plotted as a function of the arithmetic distance from the load head. The deflection at the position corresponding to the Number 1 sensor is the Number 1 projected deflection as shown in Fig. 1 which is an example difference between Number 1 sensor deflection (1*M*) and Number 1 projected deflection (1*P*) for normal pavement behavior (14,15): Number 1 projected = exp [(2 log Number 2 deflection) − (log Number 3 deflection)]. The slope of the semi-log line (secant line), the difference in magnitude between the Number 1 projected and the Number 1 sensor deflections, and the magnitude of all deflections are indicative of the shape of the deflection bowl.

For a given pavement structure, asphaltic concrete modulus, and subgrade modulus, there is a difference between the Number 1 projected and the Number

1 sensor deflections for theoretical deflections (Fig. 1). Similarly, there is also a difference between these values for field-measured deflections. Normally, the differences between the Number 1 projected deflection and the Number 1 sensor deflection for both theory and field measurements are the same. Slab deterioration is indicated when field measurements indicate a Number 1 sensor

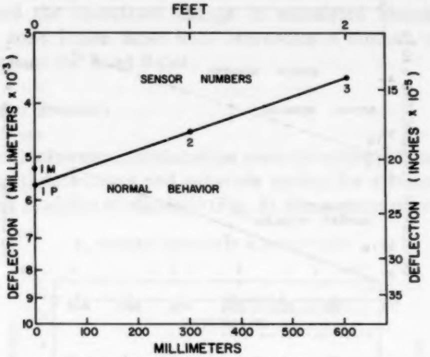


FIG. 1.—Deflection versus Distance from Load Head and Determination of Number 1 Projected Deflection for Normal Pavement Behavior (14,15)

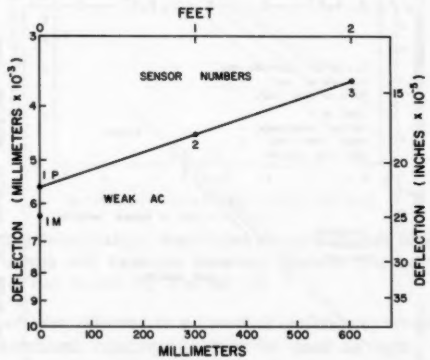


FIG. 2.—Deflection versus Distance from Load Head and Determination of Number 1 Projected Deflection for Pavement with Weak Asphaltic Concrete Layer (14,15)

deflection greater than the Number 1 projected deflection as shown in Fig. 2 which is an example difference between Number 1 sensor deflection (1*M*) and Number 1 projected deflection (1*P*) for a pavement with a weak asphaltic concrete layer (14,15). The difference between these values is greater than the difference for theoretical deflections. A foundation problem, or lack of supporting capability, may be indicated by increased magnitudes of all field deflections

and a Number 1 projected deflection greater than the Number 1 sensor deflection as shown in Fig. 3 which is an example difference between Number 1 sensor deflection (1*M*) and Number 1 projected deflection (1*P*) for a pavement with a foundation support problem (14,15). Also, the difference between the Number

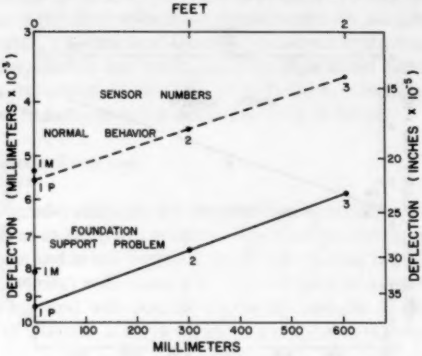


FIG. 3.—Deflection versus Distance from Load Head and Determination of Number 1 Projected Deflection for Pavement with Foundation Support Problem (14,15)

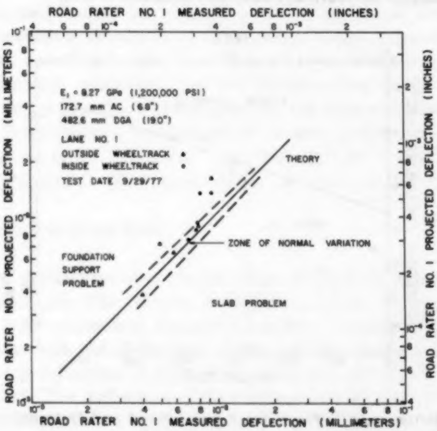


FIG. 4.—Relationship between Road Rater Number 1 Projected Deflection and Road Rater Number 1 Sensor Deflection [Field Measurements Indicate Foundation Support Problem (14,15)]

1 projected deflection and the Number 1 sensor deflection for field measurements should be greater than the difference for theoretical deflections.

A plot of Number 1 projected deflections versus Number 1 sensor deflections in log-log form may be used to identify variations in pavement structure. The

solid lines in Fig. 4 show the theoretical relationship between Number 1 projected and Number 1 sensor deflections for a constant structure and asphaltic concrete modulus. Subgrade modulus varies along the line. Points about the line represent field-measured deflections. The two dashed lines indicate the variation in position of the theoretical line due to changes in the magnitudes of the deflections by "plus or minus one unit" (2.54×10^{-4} mm or 1×10^{-5} in.) on the Road Rater meters and the associated change in calculated Number 1 projected deflection. The zone inside these lines represents a normal variation due to reading the meters of the Road Rater.

Estimating Subgrade Strength

Knowing layer thicknesses, relationships were developed (from elastic theory) between theoretical deflections and subgrade moduli for a constant (reference) asphaltic concrete modulus of elasticity (Fig. 5). For a given pavement structure,

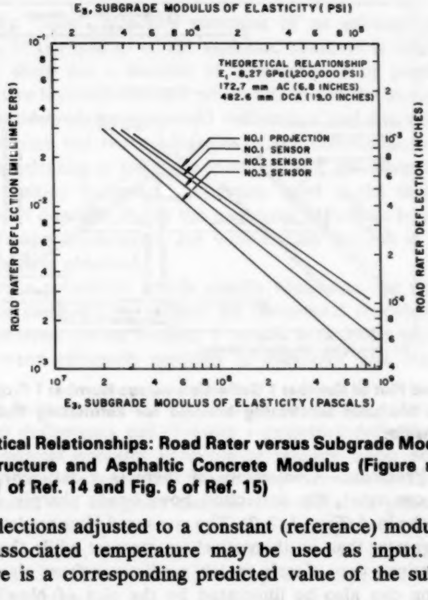


FIG. 5.—Theoretical Relationships: Road Rater versus Subgrade Modulus of Elasticity for Constant Structure and Asphaltic Concrete Modulus (Figure represents minor revision in Fig. 1 of Ref. 14 and Fig. 6 of Ref. 15)

Road Rater deflections adjusted to a constant (reference) modulus of asphaltic concrete and associated temperature may be used as input. For each field deflection, there is a corresponding predicted value of the subgrade modulus (14,15,23).

The methodology for estimating subgrade strength has evolved through several stages. Initially, the first three sensor deflections were used to obtain three estimates of the subgrade modulus. The methodology was simplified so only the Number 2 sensor deflection was used (14,15,23). Further refinements in the procedure utilize the Number 2 and Number 3 deflections to compute a Number 1 projected deflection (14,15,16,23). The Number 1 sensor deflections and Number 1 projected deflections are then plotted and compared to values predicted by elastic theory.

Interpretation of Deflection Data

Foundation or Subgrade Problems.—When a foundation or subgrade problem exists, the deflection bowl is much "broader" and "flatter" than would be theoretically expected, and the magnitudes of all the measured deflections are greater than those predicted by elastic theory (Figs. 1, 2, 3). Limited test data have indicated that excessive moisture in the subgrade could result in a measured deflection bowl of this shape (9,14,15,23). In areas where there were suspected problems with the subgrade and supporting (unbound) layers, tests indicated there was more variability among the Number 2 and Number 3 deflections than among the measured Number 1 deflections. In such a situation, either the Number 2 or Number 3 sensor deflections, or both, and the associated Number 1 projected deflections are not matching elastic theory (Figs. 3 and 6).

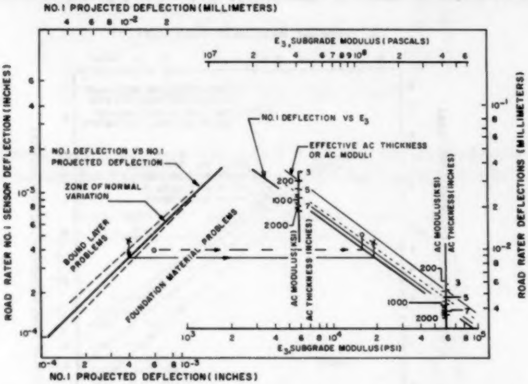


FIG. 6.—Combined Plot of Number 1 Deflection versus Number 1 Projected Deflection versus Subgrade Modulus Illustrating Method for Estimating Subgrade Strength and Effective Behavior

Bound Layer Problems.—Conversely, if there is a deficiency in the bound layer (asphaltic concrete), the deflection bowl bends sharply about the point of application of the load (Figs. 1, 2, 3). The measured Number 1 sensor deflection is considerably greater than its theoretical counterpart while the Number 2 and Number 3 deflections very closely match predictions from elastic theory (Fig. 2). This condition can also be illustrated by the plot of Number 1 projected deflection versus Number 1 sensor deflections (Fig. 6). Deflection bowls of this shape are usually observed where there are visible signs of pavement distress such as cracking and rutting (9,14,15,23).

Quantifying Effective Behavior.—Measured Road Rater deflection bowls can be evaluated using theoretical relationships. Pavement behavior (or condition) can be given in terms of a predicted subgrade modulus, effective layer thicknesses, and effective moduli of the layers. The effective behavior may be expressed as any combination of these variables which matches the measured deflections.

In this paper, pavement behavior is expressed in terms of a predicted subgrade modulus and an effective thickness of "reference" new-quality materials.

Obviously, some combinations of subgrade modulus, effective thicknesses, and layer moduli are not acceptable. An example of an unacceptable representation of pavement behavior would be an extremely low (weak) predicted value of subgrade modulus and an effective thickness of the reference material greater (thicker) than the design or as-constructed layer thicknesses. While this combination of parameters might result in a deflection bowl which resembles the measured bowl, it would not be logical for use in designing overlays because a negative calculated overlay might result. An alternative expression of pavement performance would be an increased predicted subgrade modulus and a reduced effective thickness of the reference material. Expressions of pavement behavior using asphaltic concrete moduli of elasticity other than the reference upon which the flexible pavement design curves used in Kentucky are based also would not be usable. These same curves are used in designing overlays (2,14,15,16,19, 22,23).

Determining the "true" effective structure of an existing pavement is an iterative process. If the quality of the asphaltic concrete is held fixed at some reference value, there are a number of combinations of predicted subgrade moduli and effective layer thicknesses which will match the measured deflection bowl. The magnitudes of the measured deflections and the constructed layer thicknesses determine the reasonableness of the combinations. The iterative process involves selecting a subgrade modulus and effective thicknesses and comparing the resulting theoretical deflection bowl to the measured bowl. If the theoretical bowl does not match the measured deflection bowl, the subgrade modulus and effective thicknesses are varied. This process is continued until a satisfactory match is obtained.

Figure 6 shows a procedure which usually eliminates the need for a series of iterations. The methodology utilizes the theoretical relationship of Number 1 projected deflections versus Number 1 sensor deflections and the theoretical relationship between subgrade modulus of elasticity and Number 1 sensor deflections.

In Fig. 6, the lines on the left represent a theoretical relationship between Number 1 sensor deflections and Number 1 projected deflections. The dashed lines represent normal operator variation in reading meter scales. The solid line on the right is the theoretical relationship between Road Rater Number 1 sensor deflections and subgrade moduli. Two different points are shown in Fig. 6. The "*x*'s" are data points which would be suspected of having problems in the bound layers (from the Number 1 deflection versus Number 1 projected deflection relationship). The "*o*'s" represent points suspected of having foundation of subgrade problems.

The "*x*" points (Fig. 6) have a Number 1 sensor deflection higher than would be predicted from the given values of the Number 2 and Number 3 deflections and the corresponding Number 1 projected deflections. Thus, it is necessary to adjust the deflection to match the deflections at the Number 2 and Number 3 sensors. The adjusted Number 1 deflection is then used to predict the subgrade modulus to compare to the theoretical relationship of Road Rater Number 1 deflection versus subgrade modulus. The point will plot above the theoretical line, indicating behavior weaker than the reference conditions. The behavior

may be expressed either in terms of reduced asphaltic concrete modulus or as a reduced thickness of asphaltic concrete. For overlay designs, effective behavior is more meaningful when expressed as a reduced thickness.

The "*o*" points (Fig. 6) have a Number 1 deflection lower than expected from the measured deflections at the Number 2 and Number 3 sensors and the associated Number 1 projected deflections. The deflection bowl is very "broad" and "flat" (Fig. 3) and representative of a problem in the subgrade or supporting layers. To represent the observed pavement performance in terms of a predicted subgrade modulus and effective thickness of asphaltic concrete of reference new-quality material, it is necessary to use the Number 1 sensor deflection to predict the subgrade modulus. The theoretical relationship of Number 1 projected deflection versus Number 1 sensor deflection is used in combination with the measured Number 1 deflection to determine an adjusted Number 1 deflection. The adjusted value will have a greater magnitude than the measured

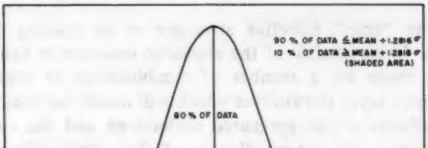


FIG. 7.—90th-Percentile Deflection

Number 1 deflection and will be compatible with the measured Number 2 and Number 3 deflections and the associated Number 1 projected deflection. When the predicted subgrade strength (based on the Number 1 sensor deflection) is plotted versus the adjusted Number 1 deflection, the expression of pavement behavior is in terms of a predicted subgrade strength and a reduced thickness of new-quality material.

Analysis of field deflections indicated this procedure will produce results which can be used as input into an overlay design process. Road Rater testing of pavements before and after overlaying shows that the ultimate behavior of the overlaid pavement is equal to that of a pavement having a total thickness of new-quality material equal to the sum of the effective thickness before overlaying and the overlay thickness (14,15,23).

Subgrade moduli may be estimated using deflections measured by any of the sensors singly or in combination, and these moduli may vary slightly. These variations usually are not significant because, for a constant asphaltic concrete modulus and any given measured deflection bowl, there is a range of subgrade moduli and asphaltic concrete thicknesses which are representative of the

measured deflections. If the measured deflection bowl and the theoretical bowl were identical, the same subgrade moduli would be predicted regardless of the way the deflections are manipulated.

Estimation of Effective Structure.—The determination of the effective pavement structure is shown by the right side of Fig. 6. If the pavement is behaving as one having a thickness equal to or greater than the theoretical or "design" thickness of asphaltic concrete, the field data will plot on the theoretical line. This also may be expressed in terms of the modulus of the asphaltic concrete—if the pavement is behaving as one having a modulus of elasticity of the asphaltic

TABLE 1.—Adjusted Road Rater Deflection Data and Predictions of Subgrade Modulus

US 60, BOYD COUNTY, KENTUCKY
EASTBOUND SHOULDER LANE
STATION 399+50 to 425+68
TEST DATE: 9/27/77
CORE THICKNESS:
172.7 mm. (6.8 in.) Asphaltic Concrete
482.6 mm. (19.0 in.) Dense Graded Aggregate
ADJUSTMENT FACTOR = 0.905
(TO ADJUST TO 70 F MEAN PAVEMENT TEMPERATURE)

ADJUSTED MEASURED DEFLECTIONS MILLIMETERS				PREDICTED SUBGRADE MODULUS MPa
No. 1	No. 2	No. 3	No. 1P	
0.00782	0.00505	0.00277	0.00927	160
0.00759	0.00528	0.00300	0.00935	167
0.00483	0.00345	0.00160	0.00739	303
0.00805	0.00691	0.00345	0.01380	155
0.00391	0.00254	0.00160	0.00396	455
0.00691	0.00460	0.00277	0.00767	186
0.00574	0.00391	0.00231	0.00665	245
0.00919	0.00691	0.00345	0.01380	128
0.00919	0.00874	0.00460	0.01660	128
0.00759	0.00620	0.00434	0.00881	167

NOTE: 1 in. = 25.4 mm
1 psi = 6,894.757 Pa

concrete equal to or stronger than that of the reference material, the field data will plot on the theoretical line. If the field data plot above the line, the pavement is performing as one made of the reference materials which is thinner than the design or theoretical thickness. Alternatively, the pavement's performance could be given in terms of a pavement of the design thickness but having a modulus of elasticity of the asphaltic concrete weaker than the reference material.

When pavement behavior is expressed in terms of reduced layer thicknesses, all layers may be varied in any combination of thicknesses of reference materials

TABLE 2.—Calculation of Effective (Behavioral) Thickness

```
US 60, BOYD COUNTY, KENTUCKY
EASTBOUND SHOULDER LANE
STATION 399+50 to 425+68
CORE THICKNESS:  172.7 mm. (6.8 in.) Asphaltic Concrete
                 482.6 mm. (19.0 in.) Dense Graded Aggregate
TEST DATE:  9/29/77

1. Read the effective structure for each data point from
   the plot of deflection versus subgrade modulus. Dense
   Graded Aggregate thickness remains constant (482.6 mm).

           EFFECTIVE ASPHALTIC CONCRETE THICKNESS
                       MILLIMETERS
                          165.1
                          152.4
                           76.2
                          121.9
                          152.4
                          172.7
                          165.1
                           88.9
                           81.3
                          167.6

     Mean                 134.4
     Standard Deviation    40.9

2. Mean Effective Structure
     134.4 mm. Asphaltic Concrete
     482.6 mm. Dense Graded Aggregate

3. Effective Structure Encompassing 90 % of the Data

     1.2816 x standard deviation = 1.2816 x 40.9 mm.
                                 = 52.4 mm. Asphaltic Concrete

   Mean - (1.2816 x standard deviation) =
     134.4 mm. - 52.4 mm. =  82.0 mm. Asphaltic Concrete

   Effective Structure
     82.0 mm. Asphaltic Concrete
     482.6 mm. Dense Graded Aggregate
```

NOTE: 1 in. = 25.4 mm.

and a predicted subgrade modulus that result in a deflection bowl which best matches the measured deflection bowl. The present procedure, however, maintains a constant crushed stone (DGA) thickness and expresses pavement behavior as a reduced thickness of asphaltic concrete at the reference modulus. If this method is used, lines of reduced thicknesses of asphaltic concrete can be superimposed onto the plot of subgrade modulus versus Number 1 sensor deflection. The effective thickness may be interpolated from these lines (Fig. 6).

Statistical analyses can be applied to either the measured Number 1 sensor deflections, the predicted subgrade moduli, or the interpolated effective thicknesses. It is recommended that any representation of pavement behavior encompass 90% of the data. Other investigators have selected similar levels (1,5,6). For example, if an effective structure is desired which encompasses 90% of

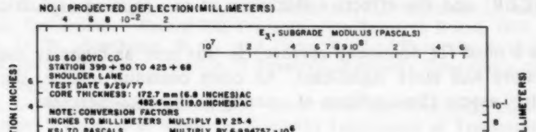


FIG. 8.—Example Analysis of Road Rater Deflections: Subgrade Moduli Predictions from Number 1 Deflections and Number 1 Projected Deflections; Graph of Number 1 Deflections versus Number 1 Projected Deflections; and Number 1 Deflections versus Subgrade Moduli

the deflection data, the recommended effective thicknesses are equal to the mean effective thickness less the product of 1.2816 and the standard deviation. Figure 7 shows the selection of the multiplier for the standard deviation. Theoretically, 90% of the measured deflections will be less than the mean deflection plus 1.2816 times the standard deviation. Note that the multiplier 1.2816 corresponds to an 80% cumulative distribution but results in a 90th percentile effective thickness because one tail of the normal distribution is not included (13,17).

Example Analysis of Road Rater Data.—Tables 1 and 2 and Fig. 8 present a set of Road Rater measurements and show the procedure to evaluate the data. This procedure utilizes concepts considered in this paper and represent modifications in earlier procedures (14,15,23).

Summary

The procedures for the evaluation of pavement performance and condition presented in this paper utilize the concepts developed and published in 1978 (14,15,23). Since then, these concepts and the associated procedure has been modified and simplified to provide a more workable procedure.

A key to an adequate design of an overlay for any pavement structure is to be able to determine reasonable values for design parameters that represent the condition of the existing pavement. The parameters considered in Kentucky's procedure include the in-place subgrade modulus and the effective thickness of the existing pavement structure at a specified reference modulus of elasticity of the asphaltic concrete. The total thickness of an overlay is equal to the difference between the thickness needed for some future design level (based on traffic volumes and associated equivalent axleloads and the in-place subgrade modulus (CBR) and the effective thickness of the existing asphaltic concrete layers.

There is a need for continued research in this area as highway maintenance becomes more and more significant. As costs continue to rise, the pressure to adequately assess the condition of existing pavements increases.

Acknowledgments

The concepts, data, and analyses reported in this paper are a result in part of Research Study KYHPR-75-77, Development of a Rational Overlay Design Method for Pavements, and Research Study KYHPR-70-49, Full-Depth Asphaltic Concrete Pavements, conducted as a part of Work Program HPR-PL-1(15), Part II, funded in part by the Federal Highway Administration and by the Kentucky Department of Transportation. The contents of the report reflect the views of the writers who are responsible for the facts and accuracy of the data presented. The contents do not necessarily reflect the official views or policies of the Kentucky Department of Transportation nor of the Federal Highway Administration. The report does not constitute a standard, specification, or regulation.

Appendix.—References

1. Bhazandes, A. C., Cumberledge, F., Hoffman, G. L., and Hopkins, J. G., "A Practical Approach to Flexible Pavement Evaluation and Rehabilitation," *Proceedings*, Fourth International Conference, Structural Design of Asphalt Pavements, University of Michigan, 1977.
2. Deen, R. C., Southgate, H. F., and Havens, J. H., "Structural Analysis of Bituminous Concrete Pavements," Division of Research, Kentucky Department of Highways, 1971.
3. Kinchen, R. W., and Temple, W. H., "Asphaltic Concrete Overlays of Rigid and Flexible Pavements," Louisiana Department of Transportation and Development, Sept., 1977.
4. Koole, R. C., "Overlay Design Based on Falling Weight Deflectometer Measurements," *Record 700*, Transportation Research Board, 1979.
5. Majidzadeh, K., "Dynamic Deflection Study for Pavement Condition Investigations," Ohio Department of Transportation, 1974.
6. Majidzadeh, K., "Pavement Condition Evaluation Utilizing Dynamic Deflection Measurements," Ohio Department of Transportation, June, 1977.

7. Michelow, J., "Analyses of Stress and Displacements in N-Layered Elastic System under a Load Uniformily Distributed on a Circular Area," Chevron Oil Research, Sept., 1963.
8. Moore, W. M., Hanson, D. I., and Hall, J. W., "An Introduction to Nondestructive Structural Evaluation of Pavements," Circular 189, Transportation Research Board, Jan., 1978.
9. "Pavement Investigation, City of Huntington Beach, CA," Project No. 10523, LaBelle Consultants, Santa Ana, Calif., 1977.
10. Peterson, G., and Shepherd, L. W., "Deflection Analysis of Flexible Pavements," Utah State Highway Department, Jan., 1972.
11. Peterson, G., "Predicting Performance of Pavements by Deflections," Utah State Highway Department, May, 1975.
12. Rufford, P. G., "A Pavement Analysis and Structural Design Procedure Based on Deflection," *Proceedings*, Fourth International Conference, Structural Design of Asphalt Pavements, University of Michigan, 1977.
13. Schwar, J. F., and Puy-Huarte, J., "Statistical Methods in Traffic Engineering," Department of Civil Engineering, The Ohio State University, Aug., 1967.
14. Sharpe, G. W., Southgate, H. F., and Deen, R. C., "Pavement Evaluation from Dynamic Deflections," *Record 700*, Transportation Research Board, 1979.
15. Sharpe, G. W., Southgate, H. F., and Deen, R. C., "Pavement Evaluation Using Road Rater Deflections," Division of Research, Kentucky Department of Transportation, Aug., 1978.
16. Sharpe, G. W., and Southgate, H. F., "Road Rater and Benkelman Beam Pavement Deflections," Division of Research, Kentucky Department of Transportation, June, 1979.
17. Snedecor, G. W., and Cochran, W. G., *Statistical Methods*, 6th ed., The Iowa State University Press, Ames, Iowa.
18. Southgate, H. F., and Deen, R. C., "Deflection Behavior of Asphaltic Concrete Pavements," Division of Research, Kentucky Department of Transportation, Jan., 1975.
19. Southgate, H. F., Deen, R. C., Havens, J. H., and Drake, W. B., "A Flexible Pavement Design and Management System," *Proceedings*, Fourth International Conference, Structural Design of Asphalt Pavements, University of Michigan, 1977.
20. Southgate, H. F., and Deen, R. C., "Temperature Distribution within Asphalt Pavements and its Relationship to Pavement Deflection," *Record 291*, Highway Research Board, 1969.
21. Southgate, H. F., and Deen, R. C., "Temperature Distribution within Asphalt Pavements," *Record 549*, Transportation Research Board, 1975.
22. Southgate, H. F., Newberry, D. C., Jr., Deen, R. C., and Havens, J. H., "Resurfacing, Restoration, and Rehabilitation of Interstate Highways: Criteria and Logic Used To Determine January 3, 1977, Needs and Estimates of Costs," Division of Research, Kentucky Department of Transportation, July, 1977.
23. Southgate, H. F., Sharpe, G. W., and Deen, R. C., "A Rational Thickness Design System for Asphaltic Concrete Overlays," *Record 700*, Transportation Research Board, 1979.
24. Timoshenko, S. P., and Goodier, J. N., *Theory of Elasticity*, 3rd ed., McGraw-Hill Book Co., Inc. New York, N.Y., 1972.
25. Vaswani, N. K., "Design of Flexible Pavements in Virginia Using AASHO Road Test Results," *Record 291*, Highway Research Board, 1970.
26. Wiseman, G., Uzan, J., Hoffman, M. S., Ishai, I., and Livneh, M., "Simple Elastic Models for Pavement Evaluation Using Measured Surface Deflection Bowls," *Proceedings*, Fourth International Conference, Structural Design of Asphalt Pavements, University of Michigan, 1977.
27. Witczak, M. W., "A Comparison of Layer Theory Design Approaches to Observed Asphalt Airfield Pavement Performance," *Proceedings*, The Association of Asphalt Paving Technologists, Vol. 44, 1975.
28. Witczak, M. W., "Asphalt Pavement Performance at Baltimore-Washington International Airport," *Research Report 74-2*, The Asphalt Institute, 1974.
29. Yoder, E. J., and Witczak, M. W., *Principles of Pavement Design*, Second ed., John Wiley and Sons, Inc., New York, N.Y., 1975.

TRANSPORTATION
ENGINEERING JOURNAL

Characterization of Asphalt Emulsion Treated Bases

By Michael S. Mamlouk[1] and Leonard E. Wood[2]

(Reviewed by the Highway Division)

Introduction

The United States highway system currently represents an enormous public investment and also has generated a large problem. New low cost, environmentally sound methods are needed to bring the system back to current design standards. The use of asphalt emulsion mixture as a paving material is a promising solution because of its ecological aspect and economic advantages. Emulsified asphalt when mixed with aggregate reduces or eliminates heating requirements. This has a significant effect on reducing energy demand and air pollution. In addition, either road mix or plant mix can be used for emulsified asphalt mix preparation. However, the most cirtical shortcoming of asphalt emulsion treated materials is the relatively low strength at early ages and the long curing time required to develop the strength which is limited by the rate of water loss from the mixture.

A thorough understanding of the integral behavior of cold mixed asphalt emulsion treated mixtures does not exist at present. Also, the resistance of the mixture to adverse moisture conditions is questionable. The purpose of this study is to characterize the cold mixed asphalt emulsion mixture used in black bases for the different mix components at different curing stages and environmental conditions. In this study, several techniques have been used to characterize the mixture. The indirect tension, resilient modulus and Hveem tests were used in the evaluation. The mixture resistance to water was also investigated. Several factors and conditions were taken into consideration.

[1]Asst. Prof., Dept. of Civ. Engrg., State Univ. of New York at Buffalo, Parker Engineering Building, Buffalo, N.Y. 14214.

[2]Prof., School of Civ. Engrg., Purdue Univ., West Lafayette, Ind. 47907.

Note.—Discussion open until August 1, 1981. To extend the closing date one month, a written request must be filed with the Manager of Technical and Professional Publications, ASCE. Manuscript was submitted for review for possible publication on February 7, 1980. This paper is part of the Transportation Engineering Journal of ASCE, Proceedings of the American Society of Civil Engineers, ©ASCE, Vol. 107, No. TE2, March, 1981. ISSN 0569-7891/81/0002-0183/$01.00.

Materials

Aggregate Types and Gradations.—Two aggregate types were used. The first type was totally a mixture of sand and gravel consisting approx of 50% calcareous and 50% siliceous pieces. About 60% of gravel particles retained on No. 4 sieve had crushed faces. The second type was totally crushed limestone. Two aggregate gradations were used in the study with a maximum size of 19 mm (3/4 in.) as shown in Fig. 1. The first was a medium gradation that followed the midspecification of the Indiana State Highway Commission (ISHC) #73B gradation band. The second was a coarse gradation which was selected at the "quarter point"; midway between the midpoint and the lower limit of the specification band. Other properties of aggregate are as shown in Table 1.

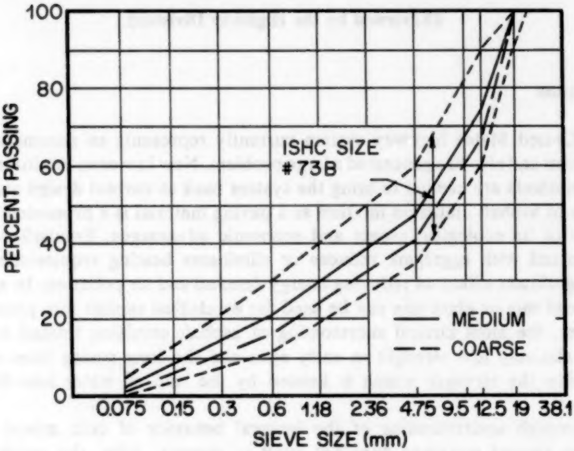


FIG. 1.—Aggregate Gradation

Asphalt Emulsion.—One high float asphalt emulsion type and grade was used which was HFMS-2s (ASTM D 977). The physical properties of the emulsion were as follows:

Saybolt Fural viscosity, *s*	50+
Residue by distillation, as a percentage	70.0
Penetration of residue after distillation, 25° C, 5s, 100 g	200+
Specific gravity of residue after distillation, 25° C	1.010

Specimen Preparation

Specimens 102 mm (4 in.) in diameter and 64 mm (2.5 in.) high were prepared according to the mix procedure suggested in a previous study (3). Different initial added moisture and asphalt emulsion contents were used. The specimens

were compacted in a fixed roller gyratory compaction machine with 1° gyration angle. A compaction effort of 20 revolutions at 1.38 MPa (200 psi) was used. The compaction process was performed at a room temperature of 22° C (72° F) with no heating provided by the compaction machine. The density was determined by measuring the height and weight of the specimen immediately after compaction.

Specimens were cured at room temperature for either 1 day or 3 days. Additional specimens were cured for 3 days in a forced-draft oven at 49° C as required by the experiment design. After curing, the specimen density was determined according to ASTM D 1188 by dusting the surface with zinc stearate instead of coating with paraffine to avoid change of moisture content (2). Three replicate specimens were prepared and tested for each combination of factors. The different tests that were used in the mixture evaluation are presented in the following paragraphs.

Indirect Tensile Test

The tensile characteristics of the asphalt emulsion mixture were determined using the indirect tensile test. The test was performed using the MTS electrohydraulic machine at temperatures of 10° C, 24° C, and 38° C (50° F, 75° F, and

TABLE 1.—Aggregate Properties

Variable (1)	Sand and gravel (2)	Limestone (3)
Apparent specific gravity	2.707	2.741
Bulk specific gravity (SSD)	2.607	2.696
Absorption, as a percentage	1.20	1.28

100° F). The load was applied at a rate of loading of 51 mm/min (2 in./min) using two curved stainless steel loading strips with a width of 12.7 mm. Continuous recordings of load versus horizontal deformation and vertical deformation versus horizontal deformation during the load application were obtained. The tensile strength, Poisson's ratio, tensile stiffness, and tensile strain at failure were also determined (5,6).

The tensile strength of the specimens is a measure of the resistance of the pavement to tensile stresses caused by traffic. High tensile strength values are required for durable asphaltic mixtures. On the other hand, high stiffness values at high temperatures are needed to reduce excessive deformation or rutting in hot weather. However, at low temperatures relatively low stiffness values are desired to eliminate or reduce cracks. The tensile strain at failure is directly related to cracking of the highway pavement. The occurrence of cracking increases as the failure strain decreases.

In this part of the study, medium graded sand and gravel and limestone mixtures were used. Two initial added moisture contents, 3% and 4.5%, and two asphalt emulsion residue contents, 3.25% and 4%, were evaluated. Both 1-day air curing and 3-day oven curing were investigated.

The tensile strengths which ranged between 25 kPa and 596 kPa were largely

affected by test temperature, aggregate type, curing, and initial added moisture content. High tensile strength values were obtained at low test temperatures, limestone mixtures, 3-day oven curing and low initial added moistures. The

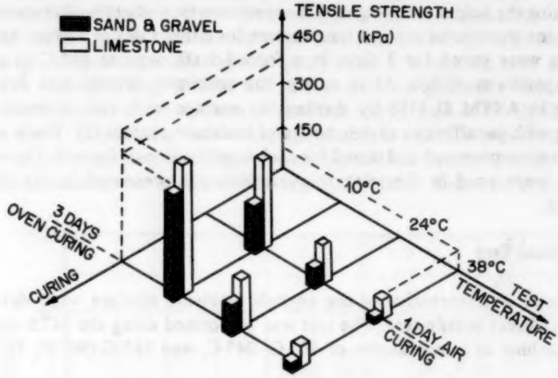


FIG. 2.—Influence of Aggregate Type, Curing, and Test Temperature on Tensile Strength

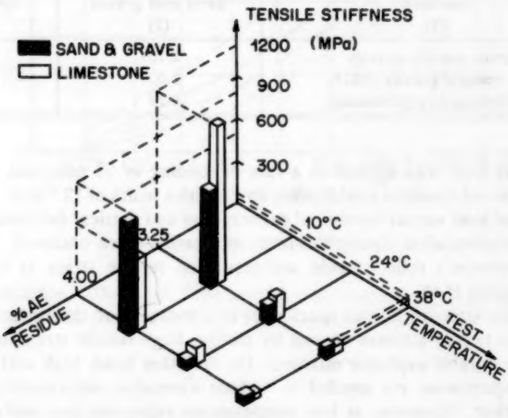


FIG. 3.—Tensile Stiffness as Function of Asphalt Emulsion Content, Test Temperature, and Aggregate Type

interaction effect of aggregate type, curing and test temperature on the tensile strength is shown in Fig. 2.

Poisson's ratio was very sensitive to the test temperature. When the test temperature increased, Poisson's ratio increased. At a test temperature of 38° C,

the specimens began to develop hair-line cracks before total failure. Therefore, Poisson's ratio values greater than 0.5 were obtained. In addition to test temperature, high values of Poisson's ratio were obtained at the 1-day air curing condition compared to the 3-day oven curing condition. In the rest of the analysis, Poisson's ratio was assumed to be 0.3, 0.35, and 0.4, at temperatures of 10° C, 24° C, and 38° C, respectively.

The influence of test temperature, aggregate type, and asphalt emulsion content, upon the tensile stiffness is presented in Fig. 3. The test temperature has an inverse effect upon the stiffness value of the mixture. It may also be noted that asphalt emulsion content and aggregate type had a marked effect upon the tensile stiffness at 10° C test temperature. Meanwhile, the tensile strain at failure was affected by the asphalt emulsion content and curing. Large tensile strains at failure were obtained for air cured mixtures with 4% asphalt emulsion residue as compared to oven cured mixtures with 3.25% asphalt emulsion residue.

Resilient Modulus Test

As traffic moves on the pavement, the vertical and horizontal stresses change such that each wheel pass can be considered as a stress pulse. The use of the dynamic tests in pavement evaluation is a realistic method because it simulates the actual stress conditions of the pavement. In this part of the study, the diametral resilient modulus test was used to evaluate the resilient characteristics of the asphalt emulsion mixes.

The resilient modulus technique used in the study was similar to the procedure developed by Schmidt (7) with some modifications (Fig. 4). A pulsating load

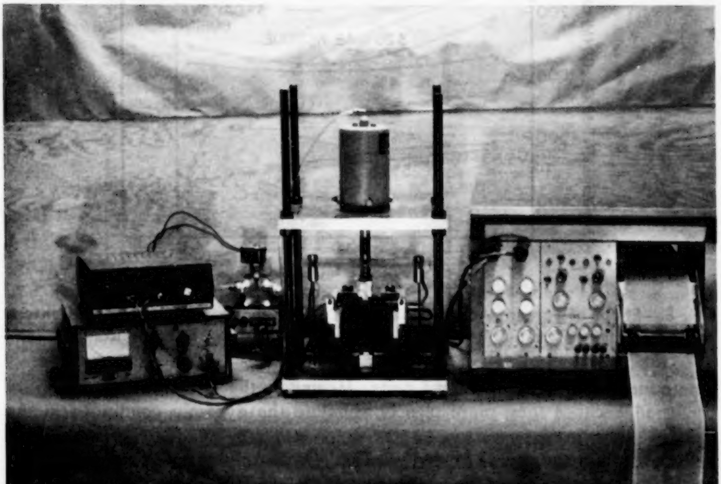


FIG. 4.—Resilient Modulus Device

of 334 N (75 lb) was applied across the vertical diameter of the specimen every 3 sec with a dwell time of 0.1 sec. Two curved stainless steel loading strips with a width of 12.7 mm were used. Both vertical and horizontal deformations were recorded on a strip chart recorder. The instantaneous resilient modulus, M_R, and the resilient Poisson's ratio were obtained according to the procedure developed by Kennedy (5). The instantaneous resilient modulus obtained from this test is analogous to the modulus of elasticity and the tensile stiffness

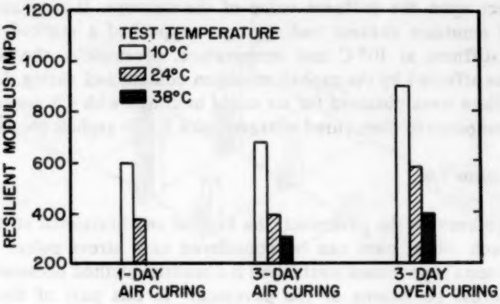


FIG. 5.—Average Instantaneous Resilient Modulus at Different Test Temperatures and Curing Levels

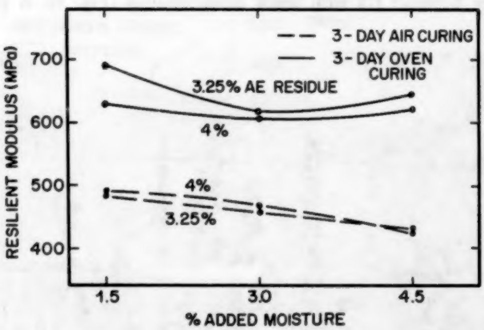


FIG. 6.—Effect of Percent Added Moisture, Curing, and Asphalt Emulsion Content on Average Instantaneous Resilient Modulus for 3-Day Air and Oven Curings

determined from the static load tests. However, the dynamic load tests are preferred because they more closely simulate the dynamic conditions in the field.

Both medium and coarse graded sand and gravel and limestone mixtures were used. Three initial added moisture contents (1.5%, 3%, and 4.5%), and two asphalt emulsion residue contents (3.25% and 4%), were investigated. Both air and oven curings were evaluated. The test was performed at temperatures of

10° C, 24° C, and 38° C. Also, since the test is nondestructive, repeated measurements were obtained from the same specimens at different temperatures in many cases.

The instantaneous resilient modulus was found to be a good measure for asphalt emulsion mixture characterization. The M_R values of the individual specimens ranged from 212 MPa to 1,157 MPa with an average of 531.9 MPa (77.2 ksi). According to the statistical results, the M_R value was largely affected

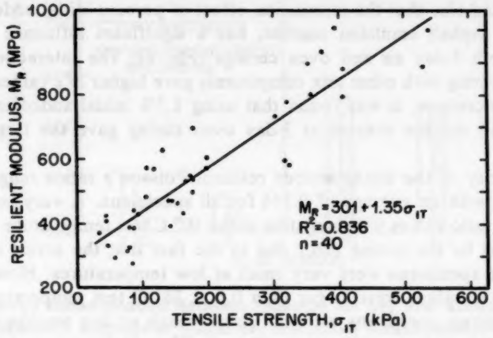


FIG. 7.—Relationship between Instantaneous Resilient Modulus and Tensile Strength

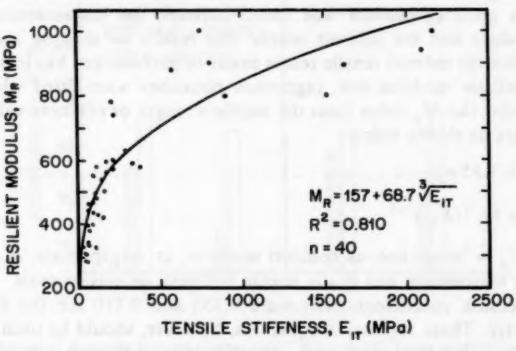


FIG. 8.—Relationship between Instantaneous Resilient Modulus and Tensile Strength

by test temperature, curing, and aggregate type. The effect of test temperature and curing on the average M_R value is demonstrated in Fig. 5. High temperatures decreased the viscosity of the asphalt residue and consequently the resilient modulus value decreased. A reduction of about 55% occurred in the average value of M_R by increasing the test temperature from 10° C–38° C for both 3-day air and oven cured specimens. On the other hand, a stiff mixture was obtained when curing was accomplished in the oven as compared to an air cured mixture.

Thus, large M_R values were obtained for oven curing relative to the air curing. Also, 3-day air cured specimens developed larger M_R values than did the 1-day air cured specimens.

The use of the limestone mixtures resulted in larger M_R values than sand and gravel mixtures. Unlike sand and gravel, the rough surface and the angular shape of the limestone as well as its mineralogical composition increased the aggregate interparticle friction and improved the coating of asphalt residue on the aggregate particles which increased the resilient modulus value.

It was found also that the interaction effect of percent initial added moisture, curing, and asphalt emulsion content, has a significant influence on the M_R value for both 3-day air and oven curings (Fig. 6). The interaction effect of 3-day oven curing with other mix components gave higher M_R values than 3-day air curing. Moreover, it was found that using 1.5% initial added moisture and 3.25% asphalt residue content at 3-day oven curing gave the largest average M_R value.

The majority of the instantaneous resilient Poisson's ratios ranged between 0.1 and 0.45 with an average of 0.314 for all specimens. A very small number of Poisson's ratio values were negative at the 10° C test temperature. The reason for this could be the testing error due to the fact that the actual deformation values of the specimens were very small at low temperatures. However, some Poisson's ratio values were larger than 0.5 at 38° C test temperature. At high test temperatures, specimens were not firm enough so that tension cracks were developed after a few applications of load. These tension cracks caused an increase in specimen volume which increased with values of Poisson's ratio.

Relationship between Resilient Modulus Test Results and Indirect Tensile Test Results.—A good correlation was found between the instantaneous resilient modulus values and the indirect tensile test results as may be seen in Figs. 7 and 8. Since the indirect tensile test is easier to perform and has less variability than the resilient modulus test, regression equations were fitted which can be used to predict the M_R value from the tensile strength or stiffness values. These equations are as shown below:

$$M_R = 301 + 1.35\sigma_{IT} \quad (1)$$

$$M_R = 157 + 68.7(E_{IT})^{1/3} \quad (2)$$

in which M_R = instantaneous resilient modulus, in megapascals; σ_{IT} = tensile strength, in kilopascals; and E_{IT} = tensile stiffness, in megapascals.

The regression coefficients, R^2, were 0.836 and 0.810 for the Eqs. 1 and 2, respectively. These regression equations, however, should be used for values of σ_{IT} up to 550 kPa and E_{IT} up to 2,000 MPa that were obtained in this study.

Water Sensitivity Test

The effect of water on the performance of the asphalt emulsion treated bases was evaluated using a modification of the Asphalt Institute water sensitivity test (8). According to this method specimens were subjected to a vacuum of 30 mm Hg for 1 h then submerged in water for 24 h at room temperature. Vacuum saturated specimens were tested using the indirect tensile test as well

as the resilient modulus test. A comparison was made between the performance of vacuum saturated specimens and the corresponding dry specimens.

A general reduction in the tensile stiffness of the vacuum saturated specimens was observed compared to dry specimens. The stiffness values of the oven cured specimens were reduced more than the air cured specimens due to vacuum saturation. The reason for this effect was mainly due to the high amount of

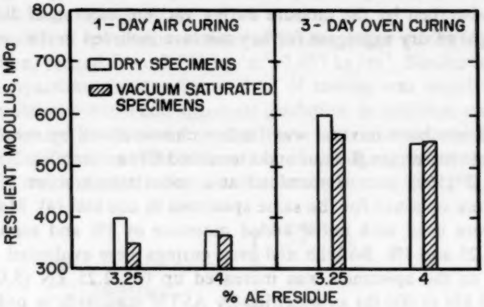


FIG. 9.—Average Instantaneous Resilient Modulus for Dry and Vacuum Saturated Specimens (Medium Aggregate Gradation, 3% Initial Added Moisture, and 24° C Test Temperature)

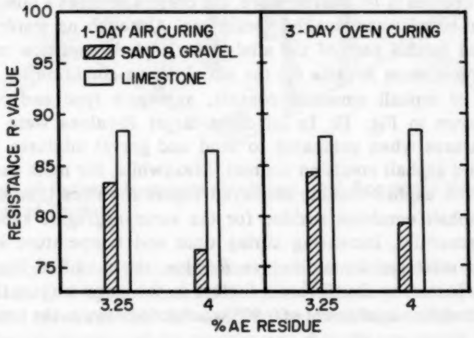


FIG. 10.—Effect of Asphalt Emulsion Content, Aggregate Type and Curing on Resistance *R*-Value (Medium Gradation and 3% Initial Added Moisture)

moisture absorbed by the oven cured specimens during vacuum saturation compared to air cured specimens. Meanwhile, the difference between the tensile strengths before and after vacuum saturation was not large and it did not show a consistent trend. In addition, the total tensile strain at failure was not largely affected by vacuum saturation.

The average resilient modulus values of dry and vacuum saturated specimens for the different asphalt emulsion contents and curing levels are shown in Fig.

9. In most cases, dry specimens gave average M_R values larger than vacuum saturated specimens. Mixtures contained 3.25% asphalt residue were affected by vacuum saturation more than mixtures contained 4% asphalt residue.

In general, the characteristics of the asphalt emulsion mixtures evaluated in this study were not severely affected by vacuum saturation. The reason could be due to the use of a densely graded mixture which had a small amount of voids and consequently small potential of moisture absorption. The amount of moisture absorbed by the mixture during vacuum saturation did not exceed 2.5% by weight of dry aggregate for any mixture included in the study.

Hveem Test

The asphalt emulsion mixture was further characterized by means of Hveem apparatus. Both resistance *R*-value and a modified Hveem stability *S*-value (ASTM D 2844 and D 1560) were determined at a room temperature of 22° C. The two values were obtained for the same specimen in one test (4). Medium graded aggregates were used with initial added moisture of 3% and asphalt emulsion contents of 3.25 and 4%. Both air and oven curings were evaluated. The vertical load applied to the specimen was increased up to 22.25 kN (5,000 lb) only, and not 26.70 kN (6,000 lb) as specified by ASTM standards in order to reduce the excessive deformation to the specimen.

It was found that the *R*-value ranged between 71.9 and 89.9 for all specimens in the study. According to the design criteria recommended by Chevron (1), the minimum *R*-value for dense graded initially cured asphalt emulsion mixtures used for base coarses is 70. Furthermore, the criteria requires a minimum *R*-value of 78 for final-cured water soaked specimens. Although no water soaking test was performed in this part of the study, the asphalt emulsion mixtures used surpassed the minimum *R*-value for the initial curing condition.

The effect of asphalt emulsion content, aggregate type and curing on the *R*-value is shown in Fig. 10. In all cases larger *R*-values were obtained for limestone mixtures when compared to sand and gravel mixtures for the same curing level and asphalt emulsion content. Meanwhile, for most cases, mixtures containing 3.25% asphalt residue displayed higher *R*-values than mixtures containing 4% asphalt emulsion residue for the same aggregate type and curing level. Also, generally, increasing curing time and temperature increased the *R*-value of the mixtures. In addition to *R*-value, the modified Hveem stability *S*-value was affected by the different factors in the same way as the resistance *R*-value. A correlation coefficient of 0.805 was found between the two parameters.

Other Asphalt Emulsion Mixture Properties

Density.—The density of the asphalt emulsion mixture at the time of compaction is a measure of compactability of the mixture. It represents the density of the pavement mixture during the field construction. High densities during construction prevent further compaction by traffic that has the effect of reducing rutting or excessive permanent deformation of the pavement. On the other hand, a certain amount of air voids in the mixture is needed to increase the curing rate and to improve the drainage in the pavement.

In this study, the density of the mixture at time of compaction varied from

2,280 kg/m^3–2,483 kg/m^3. It was largely affected by the asphalt emulsion content in the mixture. Increasing the asphalt emulsion content increases the lubrication between aggregate particles and allows more compaction to occur to the mixture, which in turn increases the density. At the same time, increasing the asphalt emulsion content to a point allows more voids to be filled with asphalt that increases the density as well. In addition to the asphalt emulsion content, medium graded aggregates provided larger densities at time of compaction compared to coarse graded aggregates.

The density of the specimens at time of testing (after curing), including the moisture portion, ranged from 2,275 kg/m^3–2,457 kg/m^3. Similar to the density at time of compaction, the density at time of testing was largely affected by the asphalt emulsion content and aggregate gradation. In addition, curing changed the density because it allowed water to leave the mixture which resulted in a reduction of the density. The interaction of curing, aggregate gradation and aggregate type had a marked effect on the density at test time (Fig. 11).

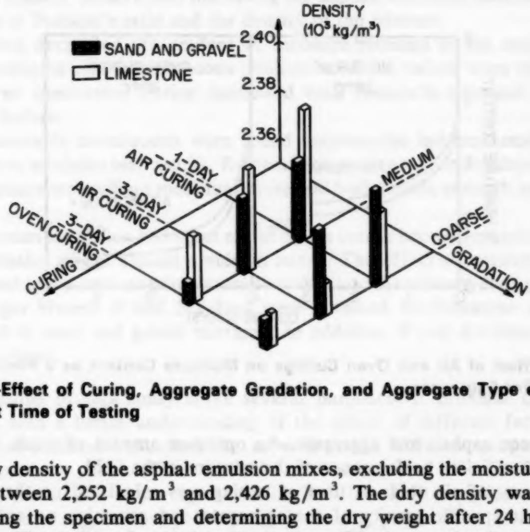


FIG. 11.—Effect of Curing, Aggregate Gradation, and Aggregate Type on Average Density at Time of Testing

The dry density of the asphalt emulsion mixes, excluding the moisture portion, varied between 2,252 kg/m^3 and 2,426 kg/m^3. The dry density was obtained by crushing the specimen and determining the dry weight after 24 h of drying at 110° C. The dry density of the mixture was directly correlated to the density after curing including the moisture portion. However, the dry density was not affected by curing. Regardless of the amount of moisture that leaves the specimen during curing, the oven dried weight of the specimen remains the same and so does the dry density.

Moisture Content.—The moisture included in the asphalt emulsion mixture comes from the water added during the initial mixture preparation as well as the moisture included in the asphalt emulsion itself. The moisture portion is very important in the preparation of the cold-mixed asphalt emulsion mixture because it increases the workability of the mix and provides a uniform coating of asphalt residue on the aggregate particles. However, a large amount of moisture

has an adverse effect on the mixture because it reduces the strength. During the curing process, the water evaporates leaving the asphalt residue adhering to the aggregate. The rate of strength development of the asphalt emulsion mixture is directly related to the rate of moisture loss from the mixture.

Fig. 12 presents the change in moisture content of the asphalt emulsion mixture for both air and oven curings at different asphalt emulsion contents. The rate of moisture lost from the mixture was large in the case of oven curing when compared to air curing. Moreover, decreasing the asphalt emulsion content increased the air voids and consequently increased the rate of moisture lost from the mixture.

Air Voids.—The amount of air voids in the asphalt emulsion bases is an influential factor that affects the curing rate of the mixture and improves the drainage in the pavement. On the other hand, increasing the air voids increases the excessive permanent deformation that could be caused by traffic (or rutting). In addition, increased air voids increases the moisture absorption in the asphalt emulsion mixes which increases the potential of stripping or weakening the

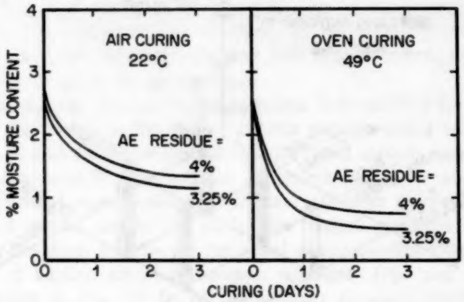


FIG. 12.—Effect of Air and Oven Curings on Moisture Content as a Percentage by Weight of Dry Aggregate

bond between asphalt and aggregate. An optimum amount of voids is needed to obtain an efficient performance of the mixture in the field.

The percent of air voids at time of testing was calculated on the bases of the apparent specific gravity of the aggregate and assuming no asphalt was absorbed into the aggregate. The amount of air voids for all specimens varied between 2.2% and 11.1% of the total volume. This wide range of percent air voids occurred because of the large effect of asphalt emulsion content, curing, aggregate type, and aggregate gradation. The percent of air voids in the asphalt emulsion mixture was closely related to the percent of total liquid at time of testing which is the sum of asphalt emulsion residue and retained moisture.

Summary and Conclusions

A comprehensive laboratory investigation was performed to characterize the cold mixed asphalt emulsion treated bases. The indirect tension, resilient modulus,

Hveem and water sensitivity tests were used in the mixture evaluation. One asphalt emulsion type, two aggregate types and gradations, two asphalt emulsion contents, and three initial added moisture contents were used. The mixture was characterized at different curing conditions. The following conclusions were obtained:

1. The asphalt emulsion mixture performance was sensitive to temperature. High test temperatures cause hair cracks to develop early during the test before complete failure for both indirect tensile and resilient modulus tests. Low tensile strengths, stiffness, and instantaneous resilient modulus values were obtained at high temperatures.
2. Limestone mixtures provided better tensile characteristics and larger instantaneous resilient modulus values than sand and gravel mixtures.
3. Mixtures containing 3.25% asphalt emulsion residue had higher tensile stiffnesses and lower tensile strains at failure than mixtures containing 4% asphalt emulsion residue. Meanwhile, increasing the asphalt emulsion content increased the value of Poisson's ratio and the density of the mixture.
4. Curing decreased the amount of moisture retained in the mixture. High tensile strengths and instantaneous resilient modulus values were obtained for oven cured specimens. Curing decreased both Poisson's ratio and the tensile strain at failure.
5. Reasonable correlations were found between the indirect tensile test and the resilient modulus test results. Regression equations were developed between the instantaneous resilient modulus values and both tensile strength and stiffness values.
6. Vacuum saturation had some effect on the tensile properties and the resilient characteristics of the asphalt emulsion mixes. The effect of vacuum saturation was related to the amount of moisture absorbed into the mixture during saturation.
7. Larger Hveem R and S values were obtained for limestone mixtures as compared to sand and gravel mixtures. In addition, R and S values were very well correlated.

The results of this study serve several purposes. It provides the highway engineer with a better understanding of the effect of different factors on the tensile and resilient characteristics as well as the different properties of the asphalt emulsion mixes. This can lead to the increase of the use of the asphalt emulsion mixture as a paving material which has a direct effect on reducing the required energy and air pollution in the paving process.

Acknowledgments

The financial support of this study from the Joint Highway Research Project, Purdue University in cooperation with the Indiana State Highway Commission (ISHC) is duly acknowledged. Sincere thanks are extended to K. E. McConnaughay, Inc., for supplying the asphalt emulsion. Appreciation is also extended to all of the people who helped in performing the tests and preparing the report.

Appendix I.—References

1. "Bitumuls Mix Manual," Chevron Asphalt Company, 1977.

2. Coyne, L. D., "Emulsion Stabilization Mix Design," presented at the 1976 Transportation Research Board Meeting.
3. Gadallah, A. A., et al., "A Suggested Method for the Preparation and Testing of Asphalt Emulsion Treated Mixtures Using Marshall Equipment," *Proceedings of the Association of Asphalt Paving Technologists,* Vol. 46, 1977, pp. 196–227.
4. Jimenez, R. A., and Morris, G. R., "Evaluating Asphalt Emulsion-Aggregate Mixtures," presented at the January, 1979, Transportation Research Board Meeting.
5. Kennedy, T. W., "Characterization of Asphalt Pavement Materials Using the Indirect Tensile Test," *Proceedings of the Association of Asphalt Paving Technologists,* Vol. 46, 1977, pp. 132–150.
6. Mamlouk, M. S., and Wood, L. E., "Evaluation of the Use of Indirect Tensile Test Results for Characterization of Asphalt Emulsion Treated Bases," *Record No. 733,* Transportation Research Board, 1979.
7. Schmidt, R. J., "A Practical Method for Measuring the Resilient Modulus of Asphalt-Treated Mixes," *Record No. 404,* Highway Research Board, 1972, pp. 22–32.
8. "Water Sensitivity Test for Compacted Bituminous Mixtures," The Asphalt Institute, The Asphalt Institute Laboratory, June, 1975.

Appendix II.—Notation

The following symbols are used in this paper:

AE = asphalt emulsion;
E_{IT} = tensile stiffness;
M_R = instantaneous resilient modulus;
n = number of specimens;
R^2 = regression coefficient;
SSD = saturated surface-dry method of determining bulk specific gravity of aggregate; and
σ_{IT} = tensile strength.

TRANSPORTATION
ENGINEERING JOURNAL

Midtown Manhattan Circulation and Surface Transit Study and Demonstration Project[a]

By Walter H. Kraft,[1] M. ASCE and Samuel I. Schwartz[2]

(Reviewed by the Urban Transportation Division)

Introduction

New York City, as the business and cultural center of the nation and the employment-entertainment center of the tri-state region, has a large demand for the movement of people and goods. The physical and economic ties between the central business district (CBD), i.e., the area of Manhattan south of 60th Street, and the remainder of the City and the region result in a focusing or concentration of the demand for transportation in the CBD. Trip generation activity in the CBD is approximately four times that of the City as a whole and three times that of the region. Transportation activity in midtown Manhattan is even more intense.

Midtown, as shown in Fig. 1, encompasses the transportation hub of the region, including Pennsylvania Railroad Station, Grand Central Railroad Station, the Port Authority Bus Terminal, and the East Side Airlines Terminal in addition to the New York City Transit Authority and the Port Authority Trans Hudson (PATH) subway systems. The grid street network is directly linked to neighboring New Jersey via the Lincoln Tunnel and to Queens and Long Island via the Queensborough (59th Street) Bridge and the Queens-Midtown Tunnel. Approximately 34% of the land area of midtown is occupied by roadways serving both trips destined to midtown locations and trips passing through midtown en route to external destinations.

Recently conducted surveys (2) indicated that 23% of all private vehicles entering midtown have destinations outside of midtown. Yet, the closing of

[a]Presented at the April 14–18, 1980, ASCE Convention and Exposition, held at Portland, Oreg. (Preprint 80-156).

[1]Vice Pres., Edwards and Kelcey, Engineering Planning Environmental Services, 1 World Trade Center, Suite 8721, New York, N.Y. 10048.

[2]Asst. Commissioner, New York City Dept. of Transportation, New York, N.Y.

Note.—Discussion open until August 1, 1981. To extend the closing date one month, a written request must be filed with the Manager of Technical and Professional Publications, ASCE. Manuscript was submitted for review for possible publication on April 17, 1980. This paper is part of the Transportation Engineering Journal of ASCE, Proceedings of the American Society of Civil Engineers, ©ASCE, Vol. 107, No. TE2, March, 1981. ISSN 0569-7891/81/0002-0197/$01.00.

the limited access West Side Highway south of 59th Street has resulted in the F.D.R. Drive on the east side being the only limited access facility to serve north-south through movements in midtown. Furthermore, there are no limited access, crosstown roadways within midtown to accommodate the demand for east-west through movement. As a result, through-trip vehicular traffic must use the midtown grid together with local traffic.

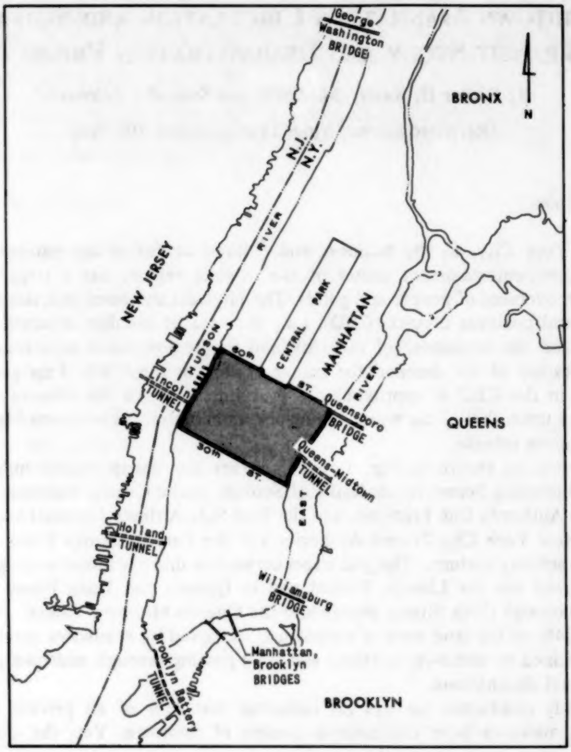


FIG. 1.—Midtown Study Area

On a typical weekday in 1976, about 2,850,000 person trips were made into the Manhattan CBD (1). Seventy percent of these were made by public transportation and 30% by auto. Approximately 700,000 person trips were made during the a.m. peak period, 90% by public transportation (bus, subway, railroad, and ferry), and the remaining 10% by automobile, taxi, or truck. It is interesting to note that although buses comprise only 3% of all motor vehicles entering the Manhattan CBD during the a.m. peak hour, they transport 48% of all persons

entering the CBD by motor vehicle and occupy only 8% of the total roadway surface occupied by all motor vehicles.

In 1977, the City of New York was mandated by the United States Court of Appeals to implement transportation strategies designated to improve air quality in midtown Manhattan by reducing vehicle usage and promoting transit. In partial response to this mandate, a project was prepared by the New York City Department of Transportation and the New York City Department of City Planning to develop and demonstrate short-range, low-cost surface transit and traffic operations improvements and to formulate these improvement techniques into an overall circulation plan and policies for midtown. The specific objectives of the project were to:

1. Improve air quality and reduce vehicular congestion through a reduction of vehicle-miles and vehicle-hours traveled, an increase in speed, a reduction in idling time, etc.
2. Increase pedestrian and vehicular safety.
3. Increase the level of surface transit service by increasing operating speeds and providing more reliable headways and improved geographical coverage.
4. Improve the integration of the surface transit system through coordination of bus routes, stops, and schedules.
5. Provide amenities for bus riders, both at the passenger boarding area as well as on-board (reduce overloading through increased frequency, improved ride quality, etc.).
6. Reduce the negative environmental impacts of buses (including traffic congestion caused by interference between buses and other modes, and environmental and visual pollution of buses in the communities traversed by bus routes and in which buses are stored).
7. Increase the efficiency of taxi operations.
8. Improve the environment for pedestrian-related activity (movement, transit boarding, window shopping, etc.).
9. Improve pedestrian flow.
10. Improve the capacity of the street system for truck operations, including movement and goods transfer.

Transportation System Model

The complex nature of surface transportation in midtown Manhattan required the development of a framework within which various elements and the interrelationship between elements could be considered. Towards this end, a conceptional model of midtown's transportation system was developed as shown in Fig. 2. The conceptual model depicts the relationship of the system to the community and the region, and of the system elements to the system. The region may be interpreted as a system consisting of numerous elements, one being the midtown community. Similarly, the midtown community may be considered as a system that contains a surface transportation system as an element. Each of the elements of a system affect that system, and the system in turn affects the larger system of which it is an element. Thus, it can be seen how an element of the surface transportation system may affect the community and in turn, the region.

The conceptual model, shown in Fig. 2, consists of physical, modal, and operational elements. The modal element includes automobiles, bicycles, buses, handcarts, mopeds, motorcycles, pedestrians, taxis, trucks, and any other mode of surface transportation. The operational element includes the motor vehicle laws of New York State, traffic regulations of New York City, posted and parking regulations, pavement marking placement and traffic signal timing. The physical element includes the roadway, sidewalks, alleys, malls, obstructions, and traffic control devices.

It has been hypothesized that all systems are comprised of three primary items: (1) Time; (2) space; and (3) information. In the foregoing conceptual model, the physical and modal elements correspond to the use of space over

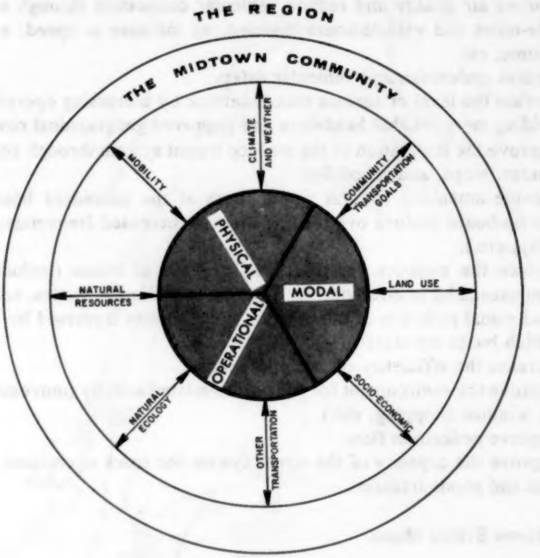


FIG. 2.—Conceptual Model of Midtown's Surface Transportation System

time, such as the length of time a vehicle occupies the roadway, while the operational element corresponds directly with information, such as the information conveyed to a driver by a traffic signal indication.

Each system exists in an environment composed of external factors that affect, or are affected by the system. For the Midtown Surface Transportation System model it is estimated that all factors can be accounted for by the eight environmental factors of mobility, national resources, national ecology, other transportation, socioeconomic, land use, community transportation goals, climate, and weather.

A system's efficiency is determined by the system's input and output. For

the Midtown Surface Transportation System, system efficiency was defined as being the ratio of person miles traveled (PMT) to person hours traveled (PHT) within the system. For example, it is apparent that the elimination of double parking will free a valuable physical element resource, i.e., roadway space, for use by a modal element unit that will, in turn, facilitate the modal element's arrival at its destination more quickly. Thus, elimination of double parking can increase "person speed" (the ratio of PMT to PHT) and increase system efficiency.

Techniques Available for Optimizing System Efficiency

Recently, much literature has been published concerning techniques to increase the efficiency of transportation in urban areas and is known as transportation system management (TSM). The focus of TSM has been to make more efficient use of existing transportation systems through low capital investment in short-range improvements. Available TSM techniques were categorized according to the elements of the Midtown Transportation System, so that a shopping list of techniques could be developed for consideration in midtown. Table 1 provides a summary of such techniques grouped according to system element primarily addressed by each technique. The efficiency-improving aspect of each technique is also summarized in the table, together with major potential impacts on system environment. For the project report each of these techniques were described including their advantages or disadvantages, or both, constraints to implementation, effectiveness, and cost of implementation and operation.

Recommended Improvement Strategies

General and specific circulation and surface transit problems in midtown were identified as a result of input from the Project's Advisory Committee, various City agencies, examination of the existing data base and direct field observations. The applicability of techniques for increasing system efficiency described previously were compared with the identified problems and specific strategies to mitigate the problems were recommended. A summary of these strategies and their impacts is contained in Table 2 and is arranged in accordance with the system element primarily affected. No strategies are recommended for the modal element, since improvements to this element would not be low cost and could most likely not be done in a short amount of time.

The recommended operational element strategies include the following:

1. Transit Operations.—The continued operation of existing and the implementation of additional exclusive bus lanes should benefit by increasing bus travel speeds and reducing intermodal conflicts. Together with the reorganization of avenue bus stops to a four-block spacing, this strategy is expected to produce a significant increase in system efficiency.

2. Traffic Operations.—Implementation of various traffic operations measures, including turn restrictions, signal timing revisions, and a traffic monitoring and response system will contribute to the establishment of a hierarchy of street use designed to expedite specific trips in midtown.

3. On-Street Parking and Loading.—Continuation of existing curb-use regula-

TABLE 1.—Summary of Techniques for Increasing System Efficiency

Primary system element (1)	Techniques (2)	Type of system efficiency increase (3)	Major impacts on system environment (4)
Operational	Transit operations techniques	Reduced bus delays	Reduced air pollution, increased efficiency for most system operators with some efficiency decreases incurred by adversely affected trips
	Traffic operations techniques	Increased operating speed	Reduced air pollution, increased efficiency for most system operators with some efficiency decreases incurred by adversely affected trips
	On-street parking and loading strategies	Increased roadway capacities	Reduced air pollution, increased efficiency for most system operators with some efficiency decreases incurred by adversely affected trips
	Enforcement and other policy techniques	Increased operating speed	Reduced air pollution, increased efficiency for most system operators with some efficiency decreases incurred by adversely affected trips
Physical	Restricted street use	Reduced delay to some modes, increased trip lengths for others	Improved environment along restricted street, potential for some VMT increases
	Line-haul techniques	Reduced delays	Improved mobility with no significant adverse impacts
	Transit techniques	Reduced delays/increased operating speed	Improved mobility with no significant adverse impacts
	Control hardware	Dependent on operations to be controlled	Improved mobility with no significant adverse impacts
Modal	Delivery vehicle modification	Increased roadway capacity	Increased efficiency benefiting all modal subelements
	Transit vehicle modification	Reduced bus delays	

TABLE 2.—Recommended Circulation Plan Strategies and Impacts

System element (1)	Strategy (2)	Type of system efficiency increase (3)	Major impacts on system environment (4)
Operational	Transit operations	Reduced bus delays resulting in increased bus travel speeds	Reduced air pollution from vehicle emissions; increased efficiency for most system operators with some efficiency decreases incurred by adversely affected trips
	Traffic operations	Increased travel speeds	Reduced air pollution from vehicle emissions; increased efficiency for most system operators with some efficiency decreases incurred by adversely affected trips
	On-street parking and loading	Increased roadway capacity potentially resulting in reduced delays and increased travel speeds	Reduced air pollution from vehicle emissions; increased efficiency for most system operators with some efficiency decreases incurred by adversely affected trips
	Enforcement and other policies	Reduced delays, through faster response to the incidents	Reduced air pollution from vehicle emissions; increased efficiency for most system operators with some efficiency decreases incurred by adversely affected trips
Physical	Restricted street use	Reduced delay to some modes; increased VMT for others; reduction in conflicts between modes	Improved environment along restricted street; potential for some VMT increases
	Transit	Reduced delays resulting in increased travel speed; reduction in conflicts between modes	Improved mobility with no significant adverse impacts
	Obstruction control	Reduced delays to all modes; improved conditions for pedestrians	Improved midtown environment

tions in midtown, modified to reduce the number of parking spaces available at critical locations and a reduction in the supply of authorized parking spaces, should result in further increases in system efficiency through the creation of residual roadway capacity.

4. Enforcement and Other Policies.—Unification of traffic enforcement efforts under the direct control of the Department of Transportation, accompanied by continued deployment of Parking Enforcement Agents (PEA's) and Traffic Control Agents (TCA's) and by towing operations is expected to complement other plan strategies in reducing the frequency and duration of delay-causing incidents in midtown.

Recommended physical element strategies include the following:

1. Restricted Street Use.—Implementation of transit streets through restrictions on the use of these streets by other modes should result in a significant increase in the travel speeds of affected crosstown buses, while increasing the vehicle miles traveled (VMT) by other modes. The creation of transit streets will also result in an improved environment for pedestrians.

2. Transit.—The creation of express and charter bus layover areas in the vicinity of their respective initial stops in midtown will result in a reduction in bus VMT and in a reduction of conflicts between these buses and other vehicles. It is anticipated that system efficiency will increase due to the resultant increase in travel speeds of all vehicles in the vicinity of initial express bus stops.

3. Obstruction Control.—Improved control over the creation of roadway obstructions, such as street openings for utility repairs and construction activities, should reduce delays to all midtown traffic. The relocation of obstacles to pedestrian movements at critical locations will also improve pedestrian circulation, while generally improving the midtown environment.

Recommended Improvement Projects

A number of recommended improvement projects were developed based upon identified problem locations and in consideration of the strategies proposed in Table 2. Operational improvement projects were divided into area-wide and locational improvements. The area-wide improvements included a traffic monitoring system, changes in the enforcement of traffic and parking regulations, and changes in authorized parking. The establishment of a traffic monitoring system was proposed, since it would provide a mechanism for determining changes in the demands placed on midtown's transportation system and the performance of the system in accommodating such demands. The importance in establishing a traffic monitoring system in the Manhattan CBD was recognized by those responsible for formulating the Transportation Control Plan (TCP) and was included as a requirement of the TCP. The mandated system will enable the assessment of the effectiveness of the TCP strategies on a periodic basis.

Enforcement of traffic and parking regulations is an integral part of the operational element of any transportation system. Inefficient enforcement results in ineffective use of the physical element by modal element operators. Efficient enforcement is an operational technique that can provide modal operators with

TABLE 3.—Summary of Recommended Improvement Projects

Recommended improvement (1)	Major elements (2)
Traffic monitoring system	Detection subsystem-90 directional roadway stations; Communication subsystem; Processing subsystem-microprocessors at 80 monitoring stations; minicomputer
Comprehensive traffic and parking enforcement policy	Consolidate enforcement responsibility; Enforcement agent deployment plan; Communications network; Public information campaign
Authorized vehicle parking control	Zero-based allocation of spaces; Joint state and city commission to monitor the issuance of license plates to authorized vehicles; Eliminate DPL/FC de jure immunity from fines; Eliminate government agency de facto immunity from fines; Tow authorized vehicles from critical locations; Not more than 50% of blockface reserved for authorized vehicles; Relocate authorized spaces situated on major roadways in an isolated manner; Short-term duration parking for NYP, MD, and government vehicles; parking meters for NYP and MD spaces; Publish list of scofflaws
Bus stop signing	Revise signing to be consistent with no stopping or no standing regulations on blockface
5th Avenue bus lane	34th-60th Streets; Buses and right turns only; Dual-lane 7–10 a.m. and 4–7 p.m.; Single-lane 10 a.m.–4 p.m.; 4-block bus stop spacing
Madison Avenue bus lane	34th-60th Streets; Buses and right turns only; 34th-42nd Street segment: dual-lane 7–10 a.m.; single-lane 10 a.m.–7 p.m.; 42nd-60th Street segment: dual-lane 7–10 a.m. and 4–7 p.m.; single-lane 10 a.m.–4 p.m.; Prohibit left turns (except trucks) onto Madison Avenue from selected eastbound crosstown streets; 4-block bus stop spacing
8th Avenue bus lane	30th Street-Columbus Circle; Buses and right turns only; Single-lane 4–7 p.m.; 4-block bus stop spacing
57th Street westbound bus lane	6th-8th Avenues; Buses only; Single-lane 7–10 a.m.
34th Street express street	Selected turn prohibitions onto and off of 34th Street;

TABLE 3.—*Continued*

(1)	(2)
42nd Street express street	Selected turn prohibitions onto and off of 42nd street; Signal timing modifications in the vicinity of Grand Central Station
57th Street express street	Selected turn prohibitions onto and off of 57th Street;
3rd Avenue between 57th and 59th Streets	Selected turn prohibitions; Signal phasing and timing modifications; Deployment of traffic control agents; Parking regulation revisions
2nd Avenue contraflow bus lane	Reversed flow on 2nd Avenue between 57th Street and the Queensborough Bridge for buses only, 4–7 p.m.; Prohibit left turns from 2nd Avenue onto 59th, 58th and 57th Streets, 4–7 p.m.; 59th Street one-way eastbound between 2nd Avenue and the overpass to the upper level of the Queensborough Bridge; Special Advisory signing
49th and 50th Transit streets	Prohibit thru traffic except buses via positive signing, 8 a.m.–7 p.m.
Express bus layover at Grand Army Plaza	Designate 12th Avenue layover area; Relocate initial bus stop; 58th Street bus lane: 9th-5th Avenues; buses and right turns only; single-lane 4–7 p.m.
Express bus layover at Herald Square	Designate layover area along 12th Avenue near 34th Street (long-range); Designate layover areas on crosstown streets near Herald Square (short-range)
Express bus layover near 3rd Avenue and 34th Street	Designate layover area under FDR Drive near 34th Street (long-range); Designate layover area along 37th Street between Queens-Midtown Tunnel entrance and exit streets (short-range); Relocate initial bus stop
Charter bus parking in west midtown	Designate charter bus parking zones in west midtown
Minimization of impacts of roadway obstructions	Modify emergency permit system; Greater involvement of police department; Computer-based information system; Special pool of traffic control agents; Purchase a tow truck equipped to handle refuse containers found in Manhattan
59th and 60th Streets express streets	Extend 59th Street Express Street operation an additional hour; Consolidate existing bus stops prohibit selected right turns onto 60th Street during a.m. peak
Lincoln Tunnel area traffic improvements	Increased television surveillance of midtown roadways; Expand variable-message and trailblazer signing systems; Tunnel operation modifications;

TABLE 3.—*Continued*

(1)	(2)
	Make tunnel operations responsive to traffic conditions on midtown roadways
Grand Central area traffic improvements	Prohibit right turns from 42nd Street onto Madison Avenue; Signal timing modifications; Make 41st Street one-way westbound between Queens-Midtown Tunnel exit street and 5th Avenue; Extend pedestrian "walk" indications; Increase bus stop capacity; Eliminate Barnes dance at 42nd Street/Vanderbilt Avenue intersection
Pedestrian safety	Improved lighting at selected locations; Install fencing at selected locations; Maintenance of pavement markings; Train traffic control agents to be more aware of pedestrian needs; Strict enforcement of traffic and parking regulations; Signal timing modifications; Institute Safety Education Program

a form of information that cannot be misunderstood. At the present time, enforcement of traffic in midtown is presently the responsibility of a number of agencies. It is therefore recommended that an overall traffic and parking enforcement policy be established for midtown that would include the consolidation of responsibility of enforcement within a single agency, the determination of an enforcement deployment plan optimizing system efficiency, implementation of a communications network that will enable timely response by enforcement personnel and the design and conduct of a public information campaign to inform system users of the importance of enforcement and the benefits accrued as a result of the enforcement effort.

At the present time, there are various government employees, members of foreign diplomatic community, members of the New York City press community and medical professionals who use official and privately owned vehicles to travel in New York City in order to conduct business during the course of their working day. The Bureau of Traffic Operations has determined that these vehicles should be allocated readily accessible parking areas near certain buildings to minimize the traffic circulation impacts caused by their on-street parking. This policy has resulted in the designation of approx 1,450 authorized parking spaces in midtown. An analysis of this policy has concluded that the number of authorized parking spaces should be reassessed, many should be relocated to further minimize their adverse impacts on traffic circulation and their utilization should be restricted to short-term rather than all-day commuter parking. It was further concluded that spaces should be allocated on a basis of "zero-based parking" where each consulate, governmental agency, media group, medical institution, etc., would be required to request and justify all on-street authorized spaces annually.

TABLE 4.—Summary of Anticipated Impacts

Recommended improvement (1)	Anticipated Impacts: Improve air quality by reducing VMT or VHT (2)	Increase pedestrian and vehicular safety (3)	Increase level of surface transit service (4)	Improve integration of surface transit system (5)	Provide amenities for bus riders (6)	Reduce negative environmental impact of buses (7)	Increase efficiency of taxi operations (8)	Improve environment for pedestrian-related activity (9)	Improve pedestrian flow (10)	Improve capacity of street system for truck operations (11)	Other (12)
Traffic monitoring system	X	X									X
Comprehensive traffic and parking enforcement policy	X	X	X				X	X	X	X	
Authorized vehicle parking control	X		X							X	
Bus stop signing			X	X							
5th Avenue bus lane	X	X	X		X	X	X				
Madison Avenue bus lane	X	X	X		X	X	X				
57th Street WB bus lane	X	X	X		X	X	X				
8th Avenue bus lane	X	X	X		X	X	X				
34th Street express street	X	X	X			X			X		
42nd Street express street	X	X	X			X			X		
57th Street express street	X	X	X			X			X		
3rd Avenue between 57th and 59th Streets	X	X						X	X		
2nd Avenue contraflow bus lane	X	X	X	X	X	X		X	X		
49th and 50th Streets transit streets	X	X	X			X	X	X	X		
Express bus layover at Grand Army Plaza	X	X	X			X		X			
Express bus layover at Herald Square	X	X	X			X		X			
Express bus layover near 3rd Avenue and 34th Street	X	X	X			X		X			
Charter bus parking in west midtown	X					X		X			
Minimization of impacts of roadway obstructions	X	X									

TABLE 4.—*Continued*

(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
59th and 60th Streets express streets	X	X	X			X			X		
Lincoln Tunnel area traffic improvements	X										
Grand Central area traffic improvements	X	X						X	X		
Pedestrian safety		X						X	X		

Other area-wide operational improvements recommended included changes in curb use regulations concerning bus stop signs and the improvement of pedestrian safety by various techniques.

Physical element improvement projects were divided into modal restriction improvements, bus layover improvements and obstruction reduction improvements. Modal restriction improvements consist of modifications to the physical element to segregate various modes through restricting on-street use, i.e., with the exception of designated truck routes, and associated regulations. Bus layover improvements consist of locating layovers of charter and express buses to minimize their impact upon traffic flow and to place them in closer proximity to their first point of pickup, thereby mitigating the need for drivers to leave from a layover area and arrive at their first stop early.

The proliferation of various types of roadway obstructions in midtown has resulted in reduced traffic-carrying capacity of the street network and in significant traffic congestion at isolated areas. Recommendations were made to reduce the number and severity of roadway obstructions causing traffic bottlenecks by modification of the present emergency permit system, use of enforcement, use of a computer-based information system, deployment of a special pool of traffic control agents, and the use of a tow truck equipped to handle the unusual types of refuse containers used in Manhattan. Recommendations were also made to maximize the efficiency of midtown's pedestrian system by evaluating the placement of easily relocated items, such as planters, benches and newsstands.

A summary of the recommended improvement projects are listed in Table 3.

Anticipated Impacts

The impact of implementation for each of the recommended improvement projects was assessed with respect to the objectives of the study and to their improvement in the overall efficiency of the transportation system. The ratio of person miles traveled (PMT) to person hours traveled (PHT) within the system as a measure of system efficiency was used to estimate how an individual improvement would affect the system's operation. A tabulation of primary impacts is contained in Table 4.

Demonstration Projects

A number of the recommended improvement projects were selected as demonstration projects to indicate their effectiveness in increasing transportation system efficiency. These projects were divided into two groups as follows:

TABLE 5.—Summary of Recommended Demonstration Projects

Project identification (1)	Transportation system element addressed (2)	Anticipated impacts (3)	Estimated cost,* in thousands of dollars (4)
Exclusive bus lanes on Madison Avenue	Operational: surface transit operations subelement	Bus travel speed increases up to 20%; potential for reduction in other vehicle travel speeds; total PHT decrease	8.1
Exclusive bus lane on 8th Avenue	Operational: surface transit operations subelement	Bus travel speed increases between 10% and 15%; no anticipated adverse impacts on other vehicle travel speeds; total PHT decrease	6.0
Reorganization of bus stops on 6th Avenue	Operational: surface transit operations subelement	Bus travel speed increases up to 25%; no impact on other vehicle travel speeds	0.7
Reorganization of bus stops on East 59th and on East 60th Streets	Operational: surface transit operations subelement	Savings of 10,000 PHT/yr and 400 VHT/yr	0.7
42nd Street express street	Operational: general traffic circulation subelement	Bus travel speed increases (not quantified); some additional VMT; total PHT decrease	2.6
60th Street express street	Operational: general traffic circulation subelement	Travel speed increases (bus and other vehicles); total PHT decrease accompanied by insignificant VMT change	0.3
Operational improvements in vicinity of Queensborough Bridge	Operational: general traffic circulation subelement	Reduced VHT and PHT with minor VMT increases, resulting in overall increase in system efficiency	2.8
Authorized parking reduction on 3rd Avenue	Operational: general traffic circulation subelement	Reduced authorized parking supply; minor travel speed increases	2.0

TABLE 5.—*Continued*

(1)	(2)	(3)	(4)
Enforcement (CCTV surveillance) in Times Square Area	Operational: enforcement subelement	Decrease in enforcement agent response time to traffic incidents; unquantified travel speed increases	0.6/5.0[b]
49th–50th Street transit streets	Physical: surface transit subelement	Significant bus travel speed increases; some VMT increase; total PHT decrease; improved pedestrian environment	4.8
Express bus layover modification at Grand Army Plaza	Physical: surface transit subelement	Reduced express bus VMT, increased travel speed (all modes); total PHT reduction	1.8

[a]Includes procurement and installation of required traffic control devices; cost of operation and enforcement not included.
[b]Two alternate schemes are outlined.

1. Early action demonstration projects that were implemented during the course of the study.
2. Additional demonstration projects to be implemented after the completion of the study.

Early action demonstration projects include the following:

1. 59th Street Express Street.—Implementation of this project consisting of turn prohibitions at selected intersections on East 59th Street and revisions in the routing of the M32 bus resulted in a substantial increase in bus and vehicle travel speeds and a corresponding reduction of 52,000 person hours traveled (PHT) on this street.
2. Priority Bus Lanes on First and Second Avenues and on 42nd and 57th Streets.—Implementation of these bus priority lanes in January, 1978 consisted of increased enforcement of the proposed curb-use regulations, i.e., prohibition of stopping on the avenues, or standing on the crosstown streets, of the right curb lane. These projects resulted in bus travel speed changes ranging from −9%–+24% and other vehicle travel speed changes ranging from −10%–+23%.
3. Priority Bus Lane on Sixth Avenue.—Implementation of a priority bus lane on Sixth Avenue between 34th and 59th Streets entailed the posting of no standing curb-use regulations and of signs prohibiting the use of the right curb lane by all but buses and right-turning vehicles; and the deployment of parking agents at frequent intervals along the avenue to enforce these regulations. This change has resulted in an increase in the speed of buses by 11%.
4. Class Two Bike Lanes on Broadway and Sixth Avenue between 9th Street and 59th Street.—Bike lanes with interruptions at critical intersections, e.g., Herald Square, were implemented adjacent to the left curb lane on these avenues

in July 1978. This change has resulted in increased safety for bicyclists with up to 200 bicyclists/h being recorded during peak periods.

Additional demonstration projects were developed together with recommended evaluation techniques and estimates of their anticipated impacts as contained in Table 5.

Conclusions

Evaluations of early action demonstration projects and analyses of other recommended strategies formulated during the conduct of the Midtown Circulation and Surface Transit Sutdy, indicate that the efficiency of midtown's surface transportation system can be increased through application of low-cost techniques. Since buses accommodate 48% of all midtown surface person miles traveled (PMT), it is also concluded that improvements to surface transit operations represent a primary method for increasing total system efficiency. Furthermore, to improve system efficiency to an optimum level given existing demand, a hierarchy of street use should be established. Strategies outlined for the project are considered an incremental step in achieving such a hierarchy of street use.

The major impact of the efficiency gains described above will be a lessening in the degrees of freedom currently available to the midtown auto driver. That is, restrictions against the use of certain street segments or lanes by private autos will result in residual capacity for use by other modes or specific trip types that should benefit the system as a whole to an extent greater than presently exists.

Acknowledgments

This paper is based on studies conducted by the New York City Department of Transportation and the New York Department of City Planning with Edwards and Kelcey as prime consultant and URS/Madigan-Praeger as subconsultants. City staff participating in the project included Samuel Schwartz, Anna Lloyd, Harvey Samuelson, Jerry Cheng and Arlene Malone. Consultant staff participating in the project included the first writer, Wesley La Baugh, Martin Taub, Thomas Gawley, Raymond Tillman and Louis Gonzales.

Appendix.—References

1. Hub-Bound Travel, Tri-State Regional Planning Commission, 1978.
2. Midtown Auto Driver Study, Crossley Surveys, Inc., prepared for the New York Department of City Planning, 1977.

TRANSPORTATION
ENGINEERING JOURNAL

Allocating Public Transit Costs Among Participants

By James W. Male,[1] A. M. ASCE, John Collura,[2] A. M. ASCE,
and Paul W. Shuldiner,[3] M. ASCE

(Reviewed by the Urban Transportation Division)

Introduction

In the planning and design of a regionwide public transportation system in rural areas it has been suggested that services be coordinated in order to avoid public confusion, reduce duplication of services, increase system efficiency, and minimize waste of resources (13,14,15,18). Emphasis has been placed on the need to coordinate services in the Federal regulations for public transportation programs in nonurbanized areas (Urban Mass Transportation Act of 1964 as amended, Section 18). The regulations state that each project submitted for funding shall "contain a description of efforts to coordinate with social service agencies . . . and all transportation providers both private and public." More recent attention has been given to such coordination in the *White House Rural Initiatives* of 1979.

This paper focuses on the subject of cost allocation as it pertains to the coordination of rural public transportation services. The allocation of costs for rural public transportation is a particularly vexatious problem for a variety of reasons that relate to this type of transportation service. The most salient of these inherent characteristics include the following:

1. Rural public transportation is generally multijurisdictional. With few exceptions, the service extends to several political jurisdictions, each of which is, quite naturally, reluctant to pay any more than what it considers to be its "fair share."
2. Rural public transportation frequently is established to serve a variety of

[1] Assoc. Prof., Dept. of Civ. Engrg., Univ. of Massachusetts, Amherst, Mass. 01003.
[2] Asst. Prof., Dept. of Civ. Engrg., Univ. of Massachusetts, Amherst, Mass. 01003.
[3] Prof., Dept. of Civ. Engrg., Univ. of Massachusetts, Amherst, Mass. 01003.

Note.—Discussion open until August 1, 1981. To extend the closing date one month, a written request must be filed with the Manager of Technical and Professional Publications, ASCE. Manuscript was submitted for review for possible publication on July 8, 1980. This paper is part of the Transportation Engineering Journal of ASCE, Proceedings of the American Society of Civil Engineers, ©ASCE, Vol. 107, No. TE2, March, 1981. ISSN 0569-7891/81/0002-0213/$01.00.

clearly identified and well-organized special transportation interests. Governmental involvement in rural public transportation stems in large measure from a growing national concern for the specialized needs of the "transportation disadvantaged" (e.g., the elderly and the physically handicapped). In order to meet these needs, a large variety of small services must be provided, many of which prove to be too small to be operated efficiently as separate entities.

3. Rural public transportation is supported financially by a broad variety of governmental agencies, most of which are concerned primarily with aspects of social welfare other than transportation. However, transportation is often an essential element of the social service programs (2,3,10,12). In addition, many governmental regulations influence the type and amount of transportation that is provided.

4. Rural public transportation is highly dependent for its financial support upon sources other than fares collected directly from individual users. As a consequence of relatively low and widely-scattered demand, service costs are high and farebox or other user-derived revenues are low.

These and other difficulties notwithstanding, the coordination of transportation services is increasingly necessitated by both economics and legislative or administrative directives. And regardless of differing requirements and constraints on towns and social service agencies, the costs of transportation services must be allocated among the participants.

The objective of this paper is to present a methodology to aid in the evaluation of different means of allocating costs. The application of the methodology will be illustrated with data from Barnstable County, Mass. In order to avoid confusion, several terms will be defined at this point. Costs can be allocated by means of any number of *procedures*. These procedures can be based on *variables* (such as vehicle-miles or participant population), and on a sampling *frequency* (the frequency with which data are collected for use in the allocation procedure). Since most allocation formulas are linear equations, all procedures in this paper will consist of linear combinations of variables.

Cost Allocation

Current methods for cost allocation are varied and often specific to a given application. The procedures in use generally have evolved through both practical and political considerations. A detailed assessment of a variety of procedures can be found in Collura et al. (6). In summary, most of the procedures are based on some measure of service, characteristics of the service area, or a combination of the two. Measures of service include variables such as passenger miles, number of trips and vehicle miles, while service area characteristics include population, elderly population, or assessed property valuation. In most cases, a linear allocation formula is used after determining an appropriate weighting coefficient for each variable. When variables representing service area characteristics are used, the necessary data (e.g., population, property valuation) are relatively time invariant and therefore requires only a periodic update. Data for the usage variables can be collected continuously, or periodic samples can be taken. Various sampling methods have been used with the frequency often being determined by the amount of money available for data collection.

In most of the cases studied by Collura et al. (6), three criteria were found to be important in choosing an allocation procedure: (1) The cost of using the procedure; (2) the equity of the results; and (3) the acceptability of the procedure to the participants.

Cost of Use.—The expense incurred in using an allocation procedure adds to the overall expense of providing the transportation service. This cost is not insignificant. For example, in Barnstable County, Mass., the cost of collecting and processing the data necessary for its cost allocation procedure was approx 5% of the operating budget. In this case the costs were probably high since they represented the first year of a demonstration project. Indications are that more advanced computer hardware and software may reduce the costs to perhaps 2% or 3% of the operating budget.

Equity.—The notion of equity is most easily put as one that attempts to have everyone contribute their fair share. However, this objective is complicated because of the many possible definitions of "fair." In many cases, a fair or equitable cost allocation is one that is perceived to be fair by most of the participants.

There are two parts to the notion of equity. The first involves selection of a principle upon which to base equity. Principles for equitable allocations could be based on, e.g., the benefits received from the service, the costs incurred in providing the service, or the ability of the user to pay for the service. Many of the cost allocation procedures that have been developed have assumed that cost allocations should be based on costs incurred (5,9), and generally require a detailed assignment of costs to specific operating parameters (4). The second part of equity involves determination of how well various allocation procedures conform to the principle. Once the principle upon which to base equity has been decided, a measure of deviation, or inequity must be applied in order to evaluate different allocation procedures.

This paper will not attempt to define the most equitable cost allocation procedure nor even propose a principle upon which to base equity. These decisions, although not easy, should be made on an individual basis by the appropriate decision makers. A more complete review of this issue can be found in Ref. 8. The review that follows will provide a means by which to compare an allocation determined by a specific procedure with an ideally equitable allocation of costs. This *ideal* allocation must be based on some principle (such as benefits received), and the participants must agree that such an allocation of the costs based on this principle is the most equitable.

Acceptability.—The acceptability criterion is very subjective and often varies among participants in a regional service. There is often a distinction between participants that are social service agencies and those that are towns. Experience has shown that one of the characteristics that makes an allocation procedure acceptable to a town is that it be easy to understand. This criterion is important when a town official must convince town members to accept a specific procedure. Social service agencies, on the other hand, are concerned about their accountability to funding agencies. Thus, they prefer procedures that provide a sufficient measure of usage. Other concepts of acceptability are possible and in general, this category will include those nonquantifiable aspects.

In the past the three criteria presented may have been used to determine an appropriate allocation procedure. In most cases, however, no logical approach

was involved in the selection process, purely subjective criteria were often used, and there was little understanding of what trade-offs were involved among the criteria when selecting a procedure.

The following section will outline a method that can be used to aid in the evaluation of alternative cost allocation procedures by determining quantitative measures for the first two criteria and displaying relative merits of the procedures.

Methodology

A quantitative evaluation of cost allocation procedures requires some means of measuring the criteria defined in the previous section. Two of the criteria, cost of use and equity can be measured quantitatively and they may be arrayed graphically in order to facilitate comparison of the characteristics of different procedures. No one procedure is "best" for all regional systems. However, based on these two criteria, some procedures can be eliminated from further consideration. Those remaining must be evaluated on the basis of more subjective criteria.

Cost of Use.—The cost of using an allocation procedure includes the number of dollars spent in collecting the appropriate data and processing the information to determine allocations. This dollar measure provides a convenient quantitative measure that is easily understood.

The cost of using an allocation procedure depends on a number of factors, of which the following two are the most important:

1. The amount of data collected.—A procedure with more variables will require more time and effort to record and process data.
2. The frequency of data collection.—Some procedures may not require continuous monitoring of use. An extreme example would be the comparison between variables such as passenger miles and the population of a town. The former might require continuous (100%) data collection, whereas the latter might need only a quarterly update. The costs of data collection and processing for the passenger-mile variable could be reduced if representative data samples were taken, rather than continuous, 100% ridership data. For example, one month might be chosen as a representative sample for the year, and this decrease in data collection and processing would substantially reduce costs.

Equity.—Determining a quantifiable measure of equity is not as straightforward as the use of dollars for cost of use, but several such measures have potential application in this evaluation framework. Rossman and Graham (16) have developed equity indices which have been modified and extended in this paper.

Three different indices of equity are examined in the following. They are based on the notion of an ideal allocation and measure deviation from the ideal. This paper will not address the question of the equity principle, and therefore selection of an ideal allocation, but rather will address measurement of deviation from the ideal. However, a means of determining an ideal allocation must be available to the user. The important issue is that the equity indices can be based on a concept considered to be ideal by the participants. A group of towns and a group of social service agencies would likely differ in their opinion as to what constitutes an equitable allocation.

Three quantitative measures are proposed as possible inequity indices:

1. The maximum absolute deviation from the ideal allocation. This index is the largest difference for all participants, between the cost allocated to the agency or town and ideal costs determined for that participant. To determine the index, the difference between allocated cost and the ideal cost for each agency or town is determined. The difference which has the largest absolute value is the index. Mathematically it is:

$$\theta_A = \max_i \{|AC_i - IC_i|\} \quad (1)$$

in which AC_i = the cost allocated to participant i; and IC_i = the cost "allocated" to participant i assuming an ideal scheme.

2. The range of the deviations from the ideal allocation. Some participants might be charged more than their ideal allocation and others less. This index is the difference between the largest overcharged and the largest undercharged. Mathematically, it is:

$$\theta_B = \max_i \{AC_i - IC_i\} - \max_i \{IC_i - AC_i\} \quad (2)$$

3. The sum of deviations from the ideal. Both of the preceding measures describe the extreme (or worst) case as the index. This index sums the deviations of all participants from the ideal allocation. Mathematically it is:

$$\theta_C = \sum_i |AC_i - IC_i| \quad (3)$$

Graphical Methodology.—The selection of the cost allocation procedure that is most appropriate for a specific region will involve many factors. Among these will be the questions of equity and cost of use. These questions will be used to establish two objectives in the determination of the most appropriate procedure to minimize: (1) Inequity; and (2) the cost of use. Unfortunately, these objectives often "compete" with each other and, in most cases, a trade-off between equity and cost of use must be determined in order to decide how much a regional transit authority is willing to "pay" for equity. In some cases, however, one procedure may be less costly and more equitable than several others. In this case the other procedures are inferior to the first.

Using the quantitative measure for cost of use, and one of the indices for inequity, a more concrete comparative analysis of various allocation procedures can be made. The graphical analysis involves the following steps:

1. Using one of the equity indices and an ideal allocation, calculate the equity index for each of the procedures being considered.
2. Determine the cost of using each of the procedures.
3. On a graph in which the axes represent inequity and cost of use, plot a point for each procedure.
4. Determine the relative merits of the various procedures by eliminating inferior ones and then applying the subjective, nonquantifiable acceptability criterion to those that remain.

An example of this approach is shown in Fig. 1, in which each procedure

is plotted on a graph with axes representing inequity and cost of use. Note that in both cases the objective is to move toward the origin.

Many of the procedures can easily be eliminated as inferior, based on the criteria of equity and cost of use. For example, in Fig. 1 the point representing procedure 7 is lower and to the left of those for procedures 2, 4, 5, 9, 12, 13, 14, and 16. This means that procedure 7 is both cheaper and more equitable than the eight procedures just listed. Although no single procedure stands out as superior, a number of procedures can be isolated as being noninferior (points 1, 6, 7, 11, 15, 17, 18). Any one of the points (procedures) in the noninferior set may be an appropriate procedure to use. Again, the decision may depend on more subjective criteria. It should be noted that as one moves from point

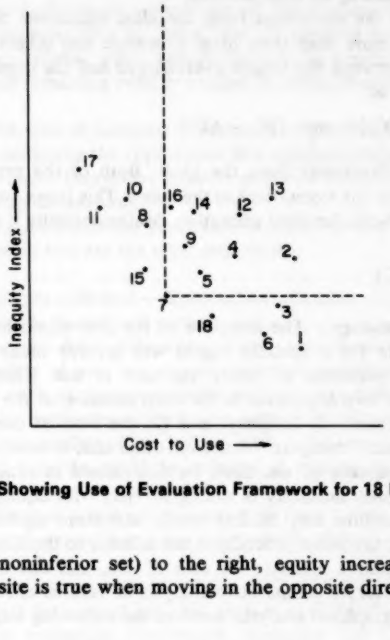


FIG. 1.—Example Showing Use of Evaluation Framework for 18 Hypothetical Procedures

to point (in the noninferior set) to the right, equity increases, as does cost of use. The opposite is true when moving in the opposite direction.

Application

In order to show the potential of the framework outlined in the previous section, data collected from Barnstable County, Mass. were used to evaluate several different allocation procedures. The procedures vary from simple to complex and represent procedures used by a variety of regional transit authorities.

Barnstable County.—Barnstable County is the governmental boundary of Cape Cod, a peninsula which extends seaward for 80 miles from the southeastern Massachusetts coastline. The Cape Cod Regional Transit Authority (CCRTA) provides demand response service to 15 member towns in Barnstable County. There are 15 towns in the county with a total population of approx 130,000.

The average density is 321 persons per mile with the primary population areas concentrated in two regions. During the summer months the population approximately triples due to the influx of summer vacationers.

The demand-response service at present provides 10,000 one-way trips every month. The CCRTA fiscal year 1978 budget provided $311,200 for operation of the service. The local share of the operating deficit was approx $50,000. This amount must be allocated among the 15 towns.

Procedures.—Table 1 shows a number of possible procedures for allocating costs among participants in a regional transit authority. The procedures range

TABLE 1.—Summary of Allocation Procedures

Procedure number (1)	Variables included (2)	Sampling frequency, as a percentage (3)	Equation (4)
1	Population	—	$AC_i = (P_i/P_T)\ AC_T$
2	Elderly population	—	$AC_i = (EP_i/EP_T)\ AC_T$
3	Property valuation	—	$AC_i = (V_i/V_T)\ AC_T$
4	Passenger trips and passenger miles	100[a]	$AC_i = (T_i/T_T)\ 0.25\ AC_T + (M_i/M_T)\ 0.75\ AC_T$
5	Passenger trips and passenger miles	100[a]	$AC_i = (T_i/T_T)\ 0.48\ AC_T + (M_i/M_T)\ 0.52\ AC_T$
6	Passenger trips and passenger miles	25[b]	$AC_i = (T_i/T_T)\ 0.25\ AC_T + (M_i/M_T)\ 0.75\ AC_T$
7	Passenger trips and passenger miles	25[b]	$AC_i = (T_i/T_T)\ 0.48\ AC_T + (M_i/M_T)\ 0.52\ AC_T$
8	Passenger trips and passenger miles	8[c]	$AC_i = (T_i/T_T)\ 0.25\ AC_T + (M_i/M_T)\ 0.75\ AC_T$
9	Passenger trips and passenger miles	8[c]	$AC_i = (T_i/T_T)\ 0.48\ AC_T + (M_i/M_T)\ 0.52\ AC_T$
10	Population, property valuation, passenger trips, and passenger miles	100[a]	$AC_i = (P_i/P_T)(AC_T/3) + (V_i/V_T)(AC_T/3) + (T_i/T_T)(AC_T/6) + (M_i/M_T)(AC_T/6)$

[a] 100% sample refers to continuous monitoring of ridership during a 12 month period.
[b] 25% sample is continuous monitoring during September, October, and November only.
[c] 8% sample is continuous monitoring during November only.

from very simple to fairly complex and involve three different structural variations: (1) Type of variable; (2) frequency of data collection; and (3) the equation combining variables (the weighting of variables). These procedures are, in some cases, variations of the currently used procedure in Barnstable County (procedure number 4 in Table 1) and, in other cases ones used by other RTAs. The 10 procedures are, by no means, a complete listing of possibilities, but a representative cross section.

Procedures 4, 6, and 8 differ from procedures 5, 7, and 9, in the coefficients applied to the variable. Procedures 4, 6, and 8 weigh passenger trips by 48%

and passenger miles by 52%. This ratio is approximately how the operating costs would be assigned to the variables if they were categorized according to fixed and variable costs respectively. The 25%/75% ratios correspond to the weightings currently used in Barnstable County. Inclusion of both weightings provides an idea of the sensitivity of the procedures to changes in the coefficient values.

Using a 100% data set from Barnstable County, each of the procedures shown in Table 1 was used to determine what cost allocations would result from their application. The data were collected with the use of a rider identification pass and daily driver logs, both of which are described in Ref. 7. For procedures using less than 100% samples, specific months from the data set were used in the calculations. The results are shown in Table 2.

TABLE 2.—Allocations Among Towns Using Procedures 1–11, in dollars

Town	Procedure 1	2	3	4	5	6	7	8	9	10	Based on individual costs
Barnstable	8,776	7,270	7,110	10,938	12,860	10,322	12,180	9,940	11,806	7,412	10,250
Bourne	6,223	2,735	3,290	3,278	1,899	2,472	2,250	2,843	2,660	2,734	2,250
Brewster	899	1,533	1,811	1,334	1,343	966	879	909	840	1,948	1,400
Chatham	3,349	4,221	4,856	917	780	824	685	633	540	3,198	850
Dennis	2,059	4,279	3,882	5,822	5,317	6,504	6,000	6,984	6,510	4,184	5,800
Eastham	1,367	1,684	2,080	630	748	643	662	848	853	1,361	600
Falmouth	9,787	7,091	6,536	8,592	7,726	8,447	7,231	7,979	6,812	7,329	8,200
Harwich	3,787	6,562	3,629	2,758	2,476	3,322	2,845	3,443	2,966	5,873	2,600
Mashpee	602	126	1,553	2,826	2,544	3,615	3,279	3,924	3,534	1,565	2,700
Orleans	848	1,957	2,512	2,694	2,993	2,504	2,775	2,576	2,897	2,164	2,500
Provincetown	1,202	885	576	4,999	5,187	3,127	3,746	2,832	3,554	2,373	4,400
Sandwich	3,612	1,630	5,317	2,049	1,868	2,459	2,195	2,304	2,105	3,091	1,800
Truro	417	100	1,130	313	263	535	438	1,046	893	432	350
Wellfleet	957	560	2,384	1,233	1,129	1,204	1,063	881	810	1,431	1,150
Yarmouth	6,117	9,367	3,342	5,040	5,304	5,466	5,750	5,282	5,643	4,969	5,150

Cost of Use.—The cost of using each of the procedures can be fairly accurately estimated, based on the costs incurred in using the current Barnstable County allocation procedure. These annual costs are approx $2,200 to make user passes and operate and maintain the necessary pass equipment, and approx $12,000 to prepare and process the data by computer. The first cost is assumed to be somewhat fixed as long as some on-board data are required, and will not vary with the type of procedure used. The second part of the total cost will vary, depending upon the amount of data processed. If data samples are processed rather than using 100% data, the data processing costs will decrease accordingly. The cost of using procedures that are not based on use (e.g., population), are assumed to have a cost of use of close to zero. The computer costs are expected to decrease with more efficient computer use.

Based on this knowledge, the cost to use each of the 10 procedures can be determined. These costs are shown in Table 3. As an example, the $5,200 cost for procedure 6 is based on the fixed cost of $2,200 (on-board data is collected), and a processing cost of $12,000/4 = $3,000, or one-fourth of the

annual costs (a 3-month sample is used).

Equity.—In order to illustrate the use of the procedure, a principle on which to base equity must be chosen. For illustration purposes the principle was to strive for allocations close to the individual costs of the participants. In other words, an equitable procedure would allocate costs in the same proportion that it would have cost the individual participants to provide the same level of service

TABLE 3.—Cost to Use and Inequity Indices for Each Procedure

Procedure number (1)	Cost of use (2)	Inequity Index θ_A (3)	Inequity Index θ_B (4)	Inequity Index θ_C (5)
1	0	3,973	7,714	25,716
2	0	4,217	7,732	26,504
3	0	4,006	7,830	27,010
4	14,200	1,028	1,138	3,849
5	14,200	2,610	3,093	6,083
6	5,200	1,273	2,188	5,876
7	5,200	1,930	2,899	6,770
8	3,200	1,568	2,792	8,576
9	3,200	1,556	2,944	9,311
10	14,200	3,273	6,111	18,081

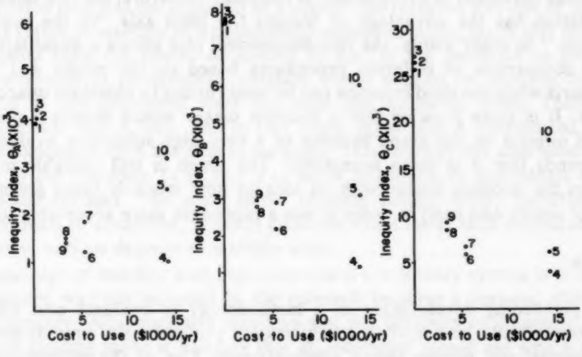


FIG. 2.—Graphical Display of Relative Merits of Ten Allocation Procedures Using Inequity Indices, θ_A, θ_B, and θ_C

if they had not been a member of the regional authority. Using the guiding principle, an ideal allocation can be determined. Once this has been done one of the three indices of inequity can be used to evaluate each of the 10 procedures.

Estimating the individual costs of the 15 towns involved the use of data on number of trips and passenger miles and representative costs for equivalent taxi rides. The use of taxi rides is based on the assumption of no group riding. This approach is not completely accurate, but is representative since most of

the RTA ridership data was demand response, door-to-door service. Using this approach individual costs were determined by multiplying the number of trips by \$1.50 per trip and multiplying the number of passenger miles by \$0.85 per passenger mile (representative taxi fares). The two results were added to get each town's individual cost. The ideal allocated costs were then determined by dividing the total cost of the regional service in proportion to the individual costs. The results are shown in the last column of Table 2.

Based on the use of the apportioned individual costs as the ideal allocation, the three equity indices (θ_A,θ_B,θ_C) were calculated. These are shown in Table 3.

Evaluation.—Based on the cost of use and equity criteria the relative merits of the 10 procedures are shown in Fig. 2 for indices θ_A, θ_B, θ_C. In all three cases a noninferior set of four procedures emerges. And in all three cases the same procedures (with one exception) form the noninferior set. Procedures 1, 4, 6, and 8 (9 for θ_A), would be chosen as having a distinct advantage over the others, but would require further consideration when compared with each other.

If a decision maker were to use the apportioned individual costs as the basis for equity and if the choice of procedure is based on cost of use and equity, the range of choice has been narrowed down to procedures 1, 4, 6, and 8 (9 if θ_A is used as the index). A choice among these four could be based on the third criterion, acceptability. This criterion can be visualized as a third axis (on a three-dimensional graph). This has not been done for two reasons: (1) A quantitative measure of acceptability is very difficult to determine; and (2) no clear definition of acceptability is possible. Therefore, the two-dimensional presentation has the advantage of leaving the third axis "in the eye of the evaluator." In other words, the two-dimensional plot allows a quantitative (and visual) comparison of different procedures based on the equity and cost of use criteria while the third criterion can be used further to eliminate unacceptable choices. It is quite possible that a decision maker would choose a procedure deemed inferior on the graph because of a very high subjective evaluation on the grounds that it is more acceptable. The graph is still valuable in that it provides the decision maker with an idea of how much is being given up (in terms of equity and cost) in order to use a procedure more acceptable to him.

Analysis

The preceeding development of a methodology to evaluate cost allocation procedures should be useful to transportation professionals, practitioners, and governmental and private sector administrators. Use of the methodology will aid decision makers by providing a quantitative means for comparison. The methodology is not intended to be used to choose one best procedure. Rather, the methodology arrays the procedures, based on certain criteria, and allows decision makers to eliminate obvious poor choices and to compare the relative merits of the remaining good possibilities. This comparison involves trade-offs among criteria as one procedure is compared with another. The process involved in evaluating the relative merits of different procedures provides valuable insight into the decision-making procedure. In addition, it allows a group of people to interact on the basis of a uniform presentation of the information.

The equity indices which were used in the analysis did not use cost per

unit service, because no universal unit of service has been agreed upon. Variables such as passenger miles, vehicle hours, population, and many more have been proposed. There may not be one acceptable unit. Additional questions involve the definition of the inequity index and how the ideal allocation of costs should be determined. This is an important question, for if decisions are made on the basis of a perceived equity (based on the value of an inequity index) then the basis of the inequity index must be sound.

Several variations in the methodology are possible which would involve changes in the calculation of values. Some of these have been emphasized as the methodology was developed. Different inequity indices can be used; three variations have been defined. In addition, any number of different procedures can be analyzed. The 10 presented are only a set of possibilities. Others could include different variables (such as vehicle miles), different sampling frequencies (such as randomly selected days in a month), and a different combination of the variables (such as a nonlinear equation).

The difference in emphasis placed on the criteria by towns and social service agencies may become a more important consideration as transportation services are further coordinated. Currently, many services operate independently, one serving specialized agencies and another serving the general public. The impetus in the future will be to reduce redundant service by combining the two. This is the current mode of operation in Barnstable County. Consolidations of this sort will increase the need to have a systematic methodology to decide on the allocation of costs since the two groups tend to emphasize the criteria differently.

A criterion that was not explored as a part of this research was one of stability. Stability refers to viability of a coordinated regional system, given that individual participants have the option of joining or not. If the incentive to join were solely economic, a participant's decision would depend on whether the costs allocated to that individual were less than the individual costs. There is no guarantee that this is true, and in fact some argue that a coordinated system is more expensive than the sum of its parts. This concept has also been carried further in the literature (1,11,16,17) to allow coalitions of individuals to form subregional systems. The most stable regional system would therefore offer all individuals or coalitions, or both, allocated costs lower than individual costs. Obviously, various degrees of stability exist.

The concept of stability is an important one if a voluntary system is to survive. The concept was not included in the research because a regional plan is not truly voluntary. Administrative directive and the difficulty in obtaining Federal and state funding make it unlikely that an individual would not join a regional plan. However, the methodology can be used for coalitions with slight modifications.

Caution should be emphasized in generalizing the results of the previous section. The analyses were presented only to illustrate the use of the methodology and not to recommend the use of certain allocation procedures. Characteristics of other transit authorities may differ yielding substantially different results. In addition, other authorities may have a unique interpretation of an ideal allocation.

Summary and Conclusions

With the increasing emphasis in rural public transportation towards consolida-

tion and coordination, the need for the allocation of costs among participants becomes obvious. The process of deciding upon a means of cost allocation is not straightforward. It involves trade-offs among equity, cost of use, and acceptability. To simplify this process a frame-work for quantitatively evaluating alternative procedures has been developed. This framework defines quantitative measures of equity and cost of use, and plots the procedures along inequity-cost axes. By comparatively ranking the procedures against inequity and cost indicators, the decision-making process is put on a more objective and less arbitrary level.

The methodology was applied using data from a demand response service in Barnstable County, Mass. Results of this application do not necessarily promote a certain cost allocation procedure but do illustrate that the methodology can provide a systematic basis for a decision.

Acknowledgments

The research reported in this paper was supported in part by the Office of Universities Research, Department of Transportation, project number DOT-RC-82028. The writers would also like to thank Lawrence Canner, Stephen Gordon, and Dale Cope for their assistance in compiling information and reviewing drafts. In addition, Robert Warren and Anthony Rogers are gratefully acknowledged for their constructive comments and assistance in gathering data.

Appendix I.—References

1. Bird, C. G., and Kortanek, K. O., "Game Theoretic Approaches to Some Air Pollution Regulation Problems," *Socio-Economic Planning Sciences,* Vol. 8, No. 3, June, 1974.
2. Brown, R. L., Lund, J., and Kidder, A. E., "Social Service Agency Transportation Study," The Transportation Institute, North Carolina Agricultural and Technical State University, Greensboro, N.C., 1973.
3. Burkhardt, J. E., Eby, C. L., Abert, J. A., Lago, A., Hedrick, J. L., and Spittel, L. A., "The Transportation Needs of the Rural Poor," *UR-072,* Resource Management Corporation, Bethesda, Md., July 1, 1969.
4. Ceglowski, K. P., Lago, A. M., and Burkhardt, J. E., "Rural Transportation Costs," *Transportation Research Record 661,* 1978.
5. Cherwony, W., and Mundle, S. R., "Transit Cost Allocation Model Development," *Transportation Engineering Journal of ASCE,* Vol. 106, No. TE1, Proc. Paper 15139, Jan., 1980, pp. 31–42.
6. Collura, J. C., Male, J. W., and Shuldiner, P. W., "Assessing Local Deficit and Social Service Agency User Charges for Rural Public Transportation," *DOT-RC-82028,* Department of Transportation, Sept., 1979.
7. Collura, J., and Warren, R. P., "Regional Paratransit Services: An Evaluation," *Transportation Engineering Journal of ASCE,* Vol. 105, No. TE6, Proc. Paper 14969, Nov., 1979, pp. 683–697.
8. Collura, J., and Cope, D. F., "Application of the Concept of Equity to Cost Allocation," *Transportation Research Record,* 1980.
9. Dierks, P. A., "Financing Urban Mass Transportation Systems: A Study of Alternative Methods to Allocate Operating Deficits," *Transportation Research Record 735,* 1979.
10. Eckmann, A., "Identifying and Serving the Elderly and Handicapped in Rural Areas," presented at the June 28, 1978, Third National Conference on Rural Public Transportation, held at Michigan Technological University, Houghton, Mich.
11. Giglio, R. J., and Wrightington, R., "Methods for Apportioning Costs Among Participants in a Regional System," *Water Resources Research,* Vol. 8, No. 5, Oct., 1972.

12. Harmon, L. J., "Service Delivery—Human Services Transportation in a Rural Region," presented at the April 5, 1977, Human Services in Rural New England Conference held at Durham, N.H.
13. "Hindrances to Coordinating Transportation of People Participating in Federally Funded Grant Programs," *Report of the Comptroller General of the United States,* Vol. 1, General Accounting Office, Oct., 1977.
14. Revis, J. S., "Social Service Transportation," *Special Report 184, Urban Transport Service Innovations,* National Academy of Sciences, Washington, D.C., 1979.
15. Rosenbloom, S., "Transportation Needs and Use of Social Services: A Reassessment," *Traffic Quarterly,* July, 1978.
16. Rossman, L. A., and Graham, P. A., "Distributing Regional Services Costs," *Journal of the Urban Planning and Development Division,* ASCE, Vol. 105, No. UP1, Proc. Paper 14284, Jan., 1979, pp. 51–82.
17. Shapley, L. S., "On Balanced Sets and Cores," *Naval Research Logistics Quarterly,* Vol. 14, No. 4, Dec., 1967.
18. Warren, R. P., Rogers, A. D., and Collura, J., "Coordination through Consolidation: The Barnstable Co. Public Transportation Program," Final Report from A National Conversation on Rural Poverty Containing Resolutions for a National Rural Policy: Formulated at Rapid City, South Dakota, Region VII CAA Association, Helena, Mont., Sept., 1979.

Appendix II.—Notation

The following symbols were used in this paper:

AC = allocated cost;
EP = elderly population;
IC = ideally allocated cost;
M = passenger-miles;
P = population;
T = passenger-trips;
V = property valuation; and
θ = equity index.

Subscripts

A, B, C = indices for different equity indices;
i = index for towns; and
T = index for total over all towns.

TRANSPORTATION ENGINEERING JOURNAL

Burial Design Criteria for Tidal Flow Crossings[a]

By Donald Steven Graham[1] and Ashish J. Mehta,[2] A. M. ASCE

(Reviewed by the Pipeline Division)

Introduction

Crossings under water bodies are among the most expensive and failure-prone sections of a pipeline. Failure accrues from many causes, including geotechnical conditions such as liquefaction and density slides, from flotation, and due to hydrodynamic forces. Burial beneath the anticipated depth of scour is the protection method of choice, but the methodology used to determine the scour depth has not received the attention it deserves. Two specific types of marine crossings have been examined here, namely tidal rivers (including estuaries), and tidal inlets. Both cases are intermediate in the continuum of watercourse-types in the sense that neither upland fluvial nor marine design criteria are strictly applicable. However, it is almost inevitable that any pipeline or cable laid along a coast will encounter several crossings of this type.

Geomorphologic Definitions

Tidal rivers are often defined to be the lower portions of rivers in which tidal fluxes are appreciable with respect to the net flow-through. A tidal river with appreciable baroclinic, i.e., density-driven, circulation is usually termed an estuary. Since the density differences are usually less than 2% locally, their presence can be ignored for the problem at hand as a first approximation, and the entire tidally-affected enclosed reach of the river may be treated *in*

[a]Presented at the April 14–18, 1980, ASCE Convention and Exposition, held at Portland, Oreg.

[1]Engr., Tudor Engineering Co., Consulting Engineers and Planners, 149 New Montgomery Street, San Francisco, Calif. 94105.

[2]Assoc. Prof., Dept. of Coastal and Oceanographic Engrg., 336 Weil Hall, Univ. of Florida, Gainesville, Fla. 32611.

Note.—Discussion open until August 1, 1981. To extend the closing date one month, a written request must be filed with the Manager of Technical and Professional Publications, ASCE. Manuscript was submitted for review for possible publication on April 22, 1980. This paper is part of the Transportation Engineering Journal of ASCE, Proceedings of the American Society of Civil Engineers, ©ASCE, Vol. 107, No. TE2, March, 1981. ISSN 0569-7891/81/0002-0227/$01.00.

toto. Enclosed tidal channels with active beds can also be included. However, a case such as a fjord that is formed by marine inundation of the coast, and in which convective transport is complicated by the presence of a sill-like bed feature near the seaward end, must be considered on a separate basis. It is in general required that the geometry of the bed of the tidal river reflect the flows occurring within it.

A tidal inlet typically is a channel or a pass through which tidal waters are exchanged between the river and the sea, and through which upland drainage passes to the sea. A stable inlet is one in which the flow through the channel is in a state of sedimentary equilibrium. This geomorphic feature is easily recognized and an analysis of appropriate burial criteria is given later.

Tidal River Flow Characteristics

It is a common misconception that the length of the "dynamic" portion of a river flowing into a "quiescent" sea is insignificant in comparison with the total length of the river. Relatively few large rivers are encountered on the coast, and most crossings are of small to intermediate size in which tidal flows dominate. Even for large rivers, the amount of water transport involved increases tremendously in the tidal reaches. For instance, it is a rule-of-thumb that each unit of freshwater entering a well-mixed U.S. East Coast estuary induces nine units of net nontidal density-current. In addition, the total tidal flux is often three to four orders of magnitude greater than the net flow-through. Accordingly, it is not surprising that both the surface and the cross-sectional areas usually increase dramatically, often exponentially, seaward of the point of tidal influence. It is also imperative to recognize that the hydrodynamics and associated geomorphic processes of tidal rivers are fundamentally different from those of upland rivers with unidirectional flows, and therefore that empirical theories developed for the latter generally are inapplicable.

The phase characteristics of flows in tidal rivers have particular relevance to pipelines. As shown in Fig. 1, the flow velocity varies with depth throughout the tidal cycle. This is true even for vertically-homogeneous flows, inasmuch as the turbulent boundary layer encompasses the entire water depth. In general the effect of density stratification is to enhance the tendency of flood flows to pass at depth and for ebb to flow out at the surface. Two points that should be noted from these observations are that: (1) Instantaneous flows can be large and in one direction (and these are the ones to design for); and (2) the tractive stress on the bottom, which causes scour, depends upon the velocity gradient near the bottom and is thus affected by density stratification and by the vertical tidal phase differences.

Design Considerations

The problem is to estimate the maximum likely scour depth over the life of the project. To accomplish this objective, some thought must be given to the design flow if it is assumed that the channel shape is determined by the bed shear (or tractive) stresses induced by the flow. The tidal river velocity $u(t)$ is often specified as the sum of several contributions according to

$$u(t) = \bar{u} + \sum_{i=1}^{i=n} \left[u_{0i} \sin\left(2\pi \frac{t}{T_i} + \phi_i\right)\right] + u'(t) \quad \ldots\ldots\ldots\ldots\ldots\ldots \text{(1)}$$

in which $\bar{u}$ = the net flow-through velocity over a tidal cycle; u_{0i}, T_i, and ϕ_i = the ith velocity amplitude, period, and phase angle of the n harmonic constituents of the flow; and $u'(t)$ represents the nonperiodic residuals. It is unlikely that the ideal design flow can be found by maximizing $\bar{u}$, u_{0i}, and u' either together or separately. Peak upland runoff often occurs in the spring and this is causally unrelated to tide or to wind. Separate tidal constituents reinforce or negate each other; consequently spring tides can have up to several times the amplitude of neap tides. The residuals are often ignored; however in regions with a low tidal range, such as the Gulf of Mexico, winds characteristically induce water level amplitudes that are two to three times that due to the astronomical tide, and therefore dominate the circulation. Hurricane-induced storm surges may very well be a design condition. Without resorting to unnecessary equations, it is simply recommended that design flows in a tidal river be based on a statistical analysis of records of $u(t)$ itself. The relative importance of the various terms of Eq. 1 can be estimated for specific cases. For instance a vernal freshet and spring tide coinciding are a likely extreme for the Fraser River in British Columbia, Canada, while a river flood or a hurricane surge (which seldom occur in the same season) would likely be appropriate to the Mississippi River. However, a concept such as one utilizing a flood of a certain year frequency based on $\bar{u}$ is probably inapplicable.

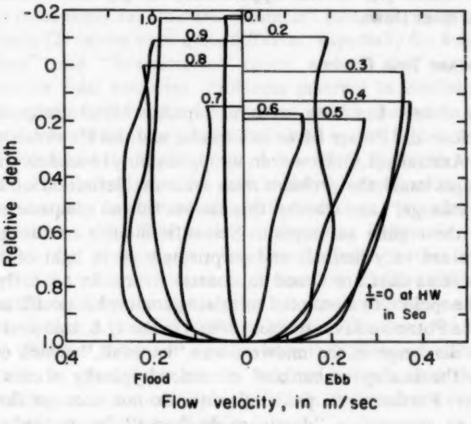


FIG. 1.—Typical Vertical Velocity Distributions over Tidal Cycle at Different Phases Represented by t/T, with Respect to High Water in Sea (Depth is Relative to Total Depth below Datum)

Regime Versus Tractive Force Theory Approaches

The two primary methodologies for the determination of river channel geometry are the regime theory and the tractive force (or rational) theory. Detailed descriptions of each can be found elsewhere (9,16,30). Briefly, the regime theory is an empirical approach that was developed from site-specific characteristics of irrigation canals in India. It has been applied to rivers, and any success it may have appears to be based on the underlying geomorphological principles (3,9,31). The tractive force theory, on the other hand, attempts to explain sediment motion, and thus bed geometry, by equating the shear stresses on the bed to the flow resistance. Again, a detailed description appears in the aforementioned references (9,16,30). The theory works best for cohesionless material with small bed forms. Neither theory was developed with oscillating flows in mind; however the tractive force theory is more general. The reader is referred to a review by the first writer (9), of the applicability to pipeline burial problems in unidirectional river flows.

Regime Theory and Tidal Rivers

The regime theory has been used for pipeline burial design in tidal rivers. Examples include the Fraser River in Canada, and the Hawkesbury River, New South Wales, Australia (2,8). However, its applicability to tidal rivers is speculative at best, since as usual the problem rests with the definition of an appropriate "dominant discharge," and whether this can provide an adequate geomorphological basis for the regime assumption. Scientific studies on the geomorphology of tidal rivers are very limited, and surprisingly so in light of the number of technical problems that are found in coastal rivers. In an early investigation Myrick and Leopold (26) attempted to relate geomorphic coefficients for a tidal tributary of the Potomac River to those found in the U.S. midwest (21). Whereas the dominant discharge in the midwest was "bankfull," which occurred every year or two, the analogous bankfull condition typically occurs semidiurnally in a tidal river. Furthermore, peak velocities do not occur at flood slack. The definition of an appropriate "dominant discharge" for unsteady flow appears to have eluded the investigators (26). For example, the definition provided is for

> flood having a maximum stage of –ft, and a range of –ft; this is one combination of range of stage and maximum stage at which maximum velocity during the tidal cycle occurs when the channel is bankfull.

This definition is clearly not based upon the tidal prism. Nevertheless they (26) found the geomorphic exponents $b, f,$ and m, to be related to the discharge, Q, according to

$$w = a\,Q^{b} \qquad (2)$$

$$h = c\,Q^{f} \qquad (3)$$

$$u = k\,Q^{m} \qquad (4)$$

in which w = the width; h = the depth; and u = the velocity. These exponents

differed significantly from those of upland rivers. In an addendum to the study, Langbein (26) showed that a closed solution yielding theoretical coefficients of $b = 0.71$, $f = 0.24$, and $m = 0.5$ could be derived by assuming: (1) An exponential decrease in the width w upstream from the mouth; (2) an exponential decrease in the depth h upstream; and (3) a uniform energy dissipation rate. The basis of the concept was originally proposed by Pillsbury (29). Partial confirmation of the underlying basis for the assumed geometry was afforded by the fact that the exponent values of b, f, and m for the tidal river studied were quite close to those derived by Langbein. A more convincing corroboration is found in the work of Harleman (13) in which the change in tidal energy dissipation for the Delaware Estuary is plotted against distance upstream and is shown to be quite uniform. The implied connection between river and estuarine regimes however is somewhat tenuous for design purposes.

Park (28) reviewed values from many sources for the geomorphic constants in Eqs. 2, 3, and 4. Relevant conclusions were that: (1) Values for "tidal estuaries" were different from those for "natural channels" (the U.S. midwest presumably having the latter); (2) values were quite different, especially for b and m, between 11 "at-a-station" and "downstream" cases; and (3) this second point was especially true for tidal estuaries. Problems inherent in confusing changes in geometry with unsteady flow at a single point versus changes in geometry with increasing flow downstream have been examined by the first writer (9).

Chantler (5) examined the applicability of the regime approach to tidal channels in more detail. He showed that the width and the depth along tidal channels do not appear to vary logarithmically with discharge, but that their product does vary logarithmically. This result is similar to O'Brien's result for tidal inlets (27). Since the product, i.e., flow area, varies to the first power, the implication is that the mean velocity remains constant. This is a different constraint from the one imposed by the assumption of constant energy dissipation. The interpretation of such a result is difficult because the flow field is oscillatory. Furthermore, Chantler does not explain how the discharge values he utilized were computed. In general his results tend to indicate that the geometry is adjusted according to some underlying principle, but that conventional regime-geomorphic relations are not applicable.

In an early paper, Blench (2) considered the application of regime theory to a hypothetical tidal river, which is apparently the Fraser River below Port Mann. It should be noted that this river is somewhat unusual in that the bedload is almost entirely sand, that the river does not widen in the usual manner near the mouth, and that the local tides have a large amplitude of approx 4 m. The basis for the application is "on the belief that, at a markedly tidal section, bed activity near dominant conditions ceases when the tide level is appreciably above its lowest value." Only hydraulic conditions at lower tide stages need then be considered. This assumption is difficult to examine or prove, but its validity might be questioned in terms of application to the case of no freshwater inflow, such as the one investigated by Myrick and Leopold (26).

Northwest Hydraulic Consultants (8) studied essentially the same case for the design of a pipeline crossing of the lower Fraser. Their approach was to calculate a design discharge on the basis of a numerical model by selecting the 100-yr return period flood at the upstream boundary and a spring tide at the lower. The regime depth, h_r, was calculated in the standard manner, namely:

$$h_r = \left(\frac{Q^2}{w^2 F_b} \right)^{1/3} \qquad (5)$$

in which F_b = the bed factor (3). The regime depth is multiplied by a z-factor to account for plan variation:

$$h_s = z\, h_r \qquad (6)$$

in which h_s = the scour depth. Finally the semi-amplitude of the dunes was added to h_s. Northwest Hydraulic Consultants concluded that

> the physical laws which govern breadth, depth and velocity in tidal estuaries have been found to differ from those of non-tidal rivers. For this reason it may be questionable to apply regime equations developed for non-tidal rivers to the design of crossings on the Fraser.

They however justified the use of Eq. 5 on the basis that high upstream river flows tended to remove tidal variations at the crossing. This would appear to contradict, to some extent, Blench's assumption (3). As noted by the first writer (9), the use of a peak discharge for the design Q in Eq. 5 would also imply that the "at-a-station" and "downstream" geomorphic coefficients were identical. As outlined by Park (28), this is not the case. However, the variations of depth with discharge appear, perhaps fortuitously, to be similar for the two cases; therefore a regime approach might work well for this parameter under selected conditions, such as characteristically occur in the Fraser River.

While the z-factors for unidirectional river flows are quite reliable (9), no comparable systematic study has been made for tidal rivers. Because the transverse and vertical velocity distributions are quite different for oscillating and stratified flows (13,33) the associated plan variations in depth can also be expected to differ. The approach used by Northwest Hydraulic Consultants may be better suited to the Fraser than to most tidal rivers, inasmuch as this river subdivides into distributaries, rather than widens, in the tidal reaches.

To summarize, Park (28) shows that the available data on the geomorphology of tidal rivers is almost negligible, even in comparison with unidirectional river flows. Furthermore, the application of regime theory is strained by the lack of an evident dominant discharge. Some preliminary results indicate that a theoretical solution may be possible on the basis of the assumption of uniform energy dissipation in the flow (26); however, the validity of this assumption is still tentative. The particular application of the empirical regime theory for scour prediction at a given location is not justifiable on the basis of the available scientific data.

Tractive Force Theory and Tidal Rivers

Because it is based solely upon the principle of force equality, the tractive force theory results should be equally valid at all locations and should be predictive. The major drawback at the present time is that solutions for three-dimensional cases are not available (9). A significant advantage of the tractive force method for pipeline burial design is that actual, rather than

"dominant," discharges can be used. These are more amenable to the design considerations as noted earlier.

For the case of a confined channel with a cohesionless bed, a simple iterative model presented by the first writer (9,10) would provide a reasonable procedure for design. For the case of hydraulically wide channels, a simpler two-dimensional analysis should apply. For tidal applications, crude estimates of the grade line at design discharges, such as have been used (9,10), would not suffice. Rather the velocity field would have to be generated by one of several available numerical models such as those of Harleman and Lee (14) for the one-dimensional case or Wang and Connor (32) for the two-dimensional case. These models assume a uniform vertical velocity, and therefore calculate a cross-sectional average or a vertically-averaged velocity for the one- and two-dimensional cases, respectively. Determination of the actual vertical velocity profile from such results is, therefore, not possible. Inasmuch as the time-scale of the tidal period is long relative to that of the vertical momentum transfer, a near-logarithmic velocity profile tends to develop during times of peak discharge which is closely related

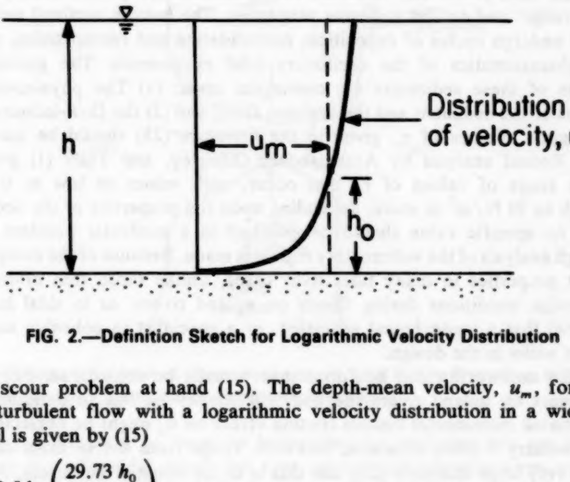


FIG. 2.—Definition Sketch for Logarithmic Velocity Distribution

to the scour problem at hand (15). The depth-mean velocity, u_m, for a fully rough turbulent flow with a logarithmic velocity distribution in a wide, open channel is given by (15)

$$\frac{u_m}{u_*} = 2.5 \ln\left(\frac{29.73\,h_0}{k_s}\right) \quad (7)$$

in which u_* = the shear velocity; k_s = Nikuradse's equivalent sand roughness; and h_0 = the elevation above the bed at which u_m equals the actual velocity, as defined in Fig. 2. By integration over the vertical it can also be shown that

$$\frac{h_0}{h} = 0.368 \quad (8)$$

Therefore $$\frac{u_m}{u_*} = 2.5 \ln\left(\frac{10.94\,h}{k_s}\right) \quad (9)$$

$$u_* = u_m \left[2.5 \ln\left(\frac{10.94\,h}{k_s} \right) \right]^{-1} \quad \ldots \quad (10)$$

Other formulations, such as one using the Manning coefficient, may be used in lieu of Eq. 10. The point is that a shear velocity can be estimated since the numerical models provide u_m and h, and k_s can be derived from measurements. By definition

$$\tau_0 = \rho u_*^2 \quad \ldots \quad (11)$$

in which τ_0 = the tractive stress; and ρ = the fluid density; τ_0 will vary with time through-out a tidal cycle, and it is emphasized that Eqs. 7, 8, 9, and 10 are not applicable during times very close to flow reversal, when the contribution from the inertia of the mass of water becomes significant relative to bed resistance (24). The next step is to find the critical tractive stress, τ_{cr}, of the bed material. For cohesionless sediments, the well-known Shields' relationship for incipient grain motion will suffice (16,25,30).

In many tidal rivers and estuaries the sediments are predominantly in the silt-clay range, and exhibit cohesive properties. The bottom surficial sediments typically undergo cycles of deposition, consolidation and resuspension, subject to the characteristics of the oscillatory tidal movements. The geotechnical properties of these sediments are contingent upon: (1) The physicochemical properties of the sediment and the ambient fluid; and (2) the flow-induced shear stress regime. Values of τ_{cr} given in the literature (23) should be used with caution. Recent analysis by Arulanandan, Gillogley, and Tully (1) points to the wide range of values of τ_{cr} that occur, with values as low as 0 N/m^2 to as high as 10 N/m^2 or more, depending upon the properties of the sediment. Clearly, no specific value should be assigned to a particular problem before a thorough analysis of the sedimentary regime is made. Because of the complicated sediment properties in many tidal river cases, unlike those that often occur under design conditions during floods on upland rivers, or in tidal inlets, it is essential that a geotechnical specialist, or a specialist in cohesive sediment dynamics assist in the design.

It is also noteworthy that bedforms must usually be considered. For design flood events on upland rivers the bed will usually be flat to antidunal, and the associated incremental bottom friction effect on u_* might be neglected (30). Typical estuary Froude numbers, however, range from 0.3 in main channels to 0.5 in very large channels (23), and this is in the range of bedforms. Accordingly, the effective, u_*, may be considerably higher. A problem is that bedform characteristics in oscillating flows are not well understood. The reversing flow will affect dune shape, and the resistance of dunes of the same height before and after flows have "turned around" will be different. Another difficulty is that the tendency for the flow to flood on one side of the channel and ebb on the other creates transversely nonhomogenous bedforms. Typical dune heights could be 1 m for sandy beds and 0.5 m or less for silt-clay ones, but this is only a very general guide. In the study reported by Northwest Hydraulic Consultants (8) for instance, 2.7-m dune amplitudes were observed. A useful reference for estimating bedform height relative to depth is due to Simons and Richardson (32). Finally, near the mouth incident waves from the sea may exert sufficient shear stresses to set the bottom in motion. Reviews for this

case have been made by Komar and Miller (19) and by Jonsson (18). Appropriate force coefficients for the case of a pipe exposed to wave action have been outlined by the first writer and Machemehl (11).

The complexity of the tidal river-estuary case has resulted in recommendations herein that are very general. These recommendations are intended to make the reader aware of some possible pitfalls, and to assist in correctly categorizing the design problem. Furthermore, the lack of scientific data upon which engineering design must be based for this type of watercourse is emphasized. Finally, it is proposed that the tractive force theory approach is more useful and applicable than one based on the regime theory.

Tidal Inlet Characteristics

Crossings at tidal inlets occur commonly, as for example in Florida, where inlets divide barrier islands that are heavily populated. Essentially, burial design

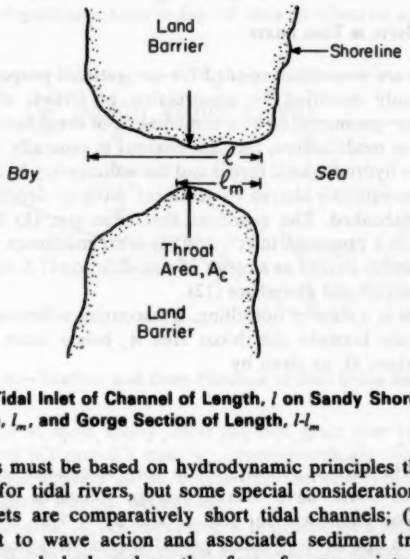


FIG. 3.—Typical Tidal Inlet of Channel of Length, l on Sandy Shoreline, with Mouth Section of Length, l_m, and Gorge Section of Length, l-l_m

criteria for inlets must be based on hydrodynamic principles that are generally similar to those for tidal rivers, but some special considerations are warranted because: (1) Inlets are comparatively short tidal channels; (2) their seaward ends are subject to wave action and associated sediment transport; and (3) most inlets have sandy beds and are therefore of common interest.

A typical tidal inlet on a sandy shoreline (Fig. 3) has a characteristic "throat," or minimum flow area whose position relative to the seaward end of the inlet channel depends upon the size of the inlet and on the wave field in the sea. In stable inlets, the throat remains more or less at the same position, but in inlets that exhibit migratory or shoaling trends the throat may move seaward or bayward to a measurable degree. The throat area also responds to flow variation in the channel, and measurements at some sandy inlets have revealed variations on the order of 10%–15% in the area about the mean value, over a tidal cycle (4). An additional factor is that even though the throat is the

smallest flow area in the inlet channel, quite often the mean depth there is at least as much as at any other section of the channel, and is sometimes greater. The throat is thus a narrow and comparatively deep section of the inlet, and is therefore selected for burial design computations.

In the seaward position of the inlet channel which is subject to wave action superimposed on the tidal flows, the bed movement can be significant, and is further influenced by storm events that themselves have a periodicity on a long-term basis. This portion of the inlet channel is often referred to as the mouth, whereas the remainder may be referred to as the gorge, as indicated in Fig. 3. The gorge is comparatively more stable than the mouth, with a much lower degree of wave-induced turbulence. The length l_m of the mouth varies; an analysis of small and intermediate-sized inlets in Florida indicated that the ratio l_m/l, in which l = the total channel length, generally varies between 0.2 and 0.5 (34). The throat usually occurs at some position in the mouth section. With reference to crossings it should be noted that inlets where they occur are usually small, and that the mouth region is the one to be avoided.

Maximum Scour Depth in Tidal Inlets

New inlet cuts are sometimes opened for navigational purposes, and existing ones are commonly modified by construction of jetties, channel dredging, alteration in the bay geometry, or by a combination of these factors. Immediately after an opening or modification, the new channel is generally "out of balance" with respect to the hydrodynamic forces and the sedimentary budget. The channel bathymetry is consequently altered by sediment scour or deposition until a new equilibrium is established. The questions that arise are: (1) What will be the maximum depth in a proposed inlet?; and (2) what maximum depth will result when the tidal prism is altered as a result of a modification? Some considerations are given in the sequel and elsewhere (12).

For sandy inlets in a state of nonsilting, nonscouring sedimentary equilibrium, a relationship exists between the throat area A_c below mean water level and the spring tidal prism, Ω, as given by

$$A_c = a_1 \Omega^{m_1} \quad \text{(12)}$$

$$\text{or} \quad \Omega = a_1^{-1/m_1} A_c^{1/m_1} \quad \text{(13)}$$

in which the coefficients a_1 and m_1, which vary somewhat from inlet to inlet and from coast to coast, have overall average values of $a_1 = 6.56 \times 10^{-5}$ and $m_1 = 1.00$ for inlets without jetties; and $a_1 = 9.01 \times 10^{-4}$ and $m_1 = 0.85$ for inlets with one or two jetties, provided Ω is measured in cubic meters and A_c in square meters. Eq. 12 was originally obtained by O'Brien (24) from observations at sandy inlets, and is applicable primarily to tides with a semidiurnal period. Jarrett (17) later compiled data on a large number of inlets and gave values of a_1 and m_1 to be used for inlets with or without jetties and located on the U.S. Pacific, Gulf or the Atlantic Coasts. Several attempts have been made to give a rational explanation for the observed dependence of Ω on A_c. Mason (22) interpreted the flow through an inlet in terms of an equivalent unidirectional flow in alluvial channels and showed that Eq. 13 agrees with the regime equations of Simons and Albertson (16). It can also be shown that the form of Eq. 13 is obtained if the Einstein-Brown bed load transport equation

(16) is interpreted in terms of an oscillatory flow. Using semianalytical considerations, Krishnamurthy (20) showed that

$$\Omega = \left[1.25 u_{*cr} T \left(1 + \frac{2a_t}{\pi h_c} \right) \ln \left(\frac{10.93 h_c}{k_s} \right) \right] A_c \quad \text{(14)}$$

in which u_{*cr} = the critical shear velocity for erosion; T = the tidal period; a_t = the inlet tidal amplitude; and h_c = the depth of the throat below mean water level. The form of Eq. 14 is similar to Eq. 13 and $m_1 = 1.00$, indicating the applicability of Eq. 14 to inlets without jetties. Furthermore, noting that T must be considered to be the semidiurnal tidal period and assuming $2a_t/\pi h_c << 1$ which is valid for typical real inlets, it may be inferred that

$$a_1 = a_1\left(u_{*cr}, \frac{k_s}{h_c} \right) \quad \text{(15)}$$

or in other words, a_1 depends on u_{*cr} and k_s/h_c only. The relative bed roughness k_s/h_c occurs as a logarithmic term in Eq. 14; thus its effect on a_1 is comparatively

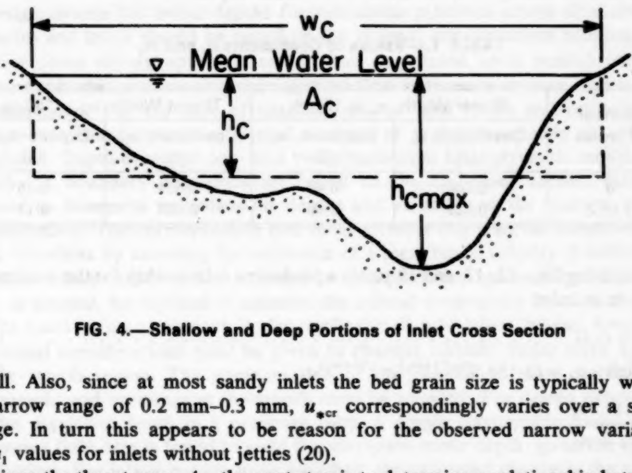


FIG. 4.—Shallow and Deep Portions of Inlet Cross Section

small. Also, since at most sandy inlets the bed grain size is typically within a narrow range of 0.2 mm–0.3 mm, u_{*cr} correspondingly varies over a small range. In turn this appears to be reason for the observed narrow variation of a_1 values for inlets without jetties (20).

Given the throat area, A_c, the next step is to determine a relationship between the surface width and the mean depth at the throat. Using regime equations for alluvial channels, Escoffier (7) utilized the relationship

$$A_c = A'_c \left(\frac{h_c}{h'_c} \right)^{5/2} \quad \text{(16)}$$

in which the prime quantities represent initial conditions. More appropriately, a large amount of data has been collected from real and model inlets (4,34). Based on these data, the following relationship between the surface width, w_c, and the mean depth, h_c, both below mean water level, such that $A_c = h_c w_c$ (Fig. 4), is obtained

$$h_c = a_2 w_c^{m_2} \quad (17)$$

in which $a_2 = 0.038$ and $m_2 = 0.87$, when $w_c \leq 150$ m; and $a_2 = 1.164$ and $m_2 = 0.19$, when $w_c > 150$ m. These values are applicable to inlets without jetties. The corresponding values for inlet with jetties are $a_2 = 0.082$, $m_2 = 0.80$ and $a_2 = 1.661$, $m_2 = 0.20$ respectively.

Fig. 4 shows that a typical throat section consists of a shallow area and a deep channel of depth, h_{cmax}. The position of this channel in the section varies from inlet to inlet, and furthermore it shifts its position anywhere between the two banks of the section during the course of time. No mathematical formulation is currently available to estimate h_{cmax}, although an analytic proof of the existence of a deep channel in a natural inlet, if such a proof can be stated, must ultimately be based on sediment transport considerations. Measurements, excluding model inlets and without differentiating between inlets with or without jetties, seem to give the relationship

$$h_{cmax} = a_3 h_c \quad (18)$$

in which an average value of $a_3 = 1.82$, thus indicating that the deep channel depth can be nearly twice as much as the mean depth.

TABLE 1.—Values of Coefficients a_4 and m_4

Number of jetties (1)	Throat Width, $w_c \leq 150$ m		Throat Width, $w_c > 150$ m	
	Coefficient, a_4 (2)	Exponent, m_4 (3)	Coefficient, a_4 (4)	Exponent, m_4 (5)
0	0.004	0.47	0.454	0.16
1 or 2	0.021	0.40	0.884	0.14

Combining Eqs. 12, 17, and 18 yields a predictive relationship for the maximum depth in an inlet

$$h_{cmax} = a_4 \Omega^{m_4} \quad (19)$$

in which $a_4 = (1.15 a_1)^{m_2/(1+m_2)} a_2^{\,1/(1+m_2)} a_3$ (20)

$$m_4 = \frac{m_1 m_2}{1 + m_2} \quad (21)$$

Here, consideration has been given to the observation that, at a given phase of the tidal cycle, the throat may be as much as 15% larger than its mean value. A quantity of interest is the relative change $\Delta h_{cmax}/h_{cmax}$ in the maximum depth for a corresponding change, $\Delta\Omega/\Omega$, in the tidal prism. Eq. 19 gives

$$\frac{\Delta h_{cmax}}{h_{cmax}} = m_4 \frac{\Delta \Omega}{\Omega} \quad (22)$$

Table 1 gives applicable values of a_4 and m_4. Comparing the m_4 values it is noted that, according to Eq. 22, the percent change in h_{cmax} for a given change in the tidal prism, Ω, is substantially more significant for small inlets ($w_c \leq 150$ m) than for larger inlets. Thus, for example, a 10% change in Ω

would result in a 47% change in h_{cmax} for a small natural inlet, but only 16% for a large inlet. The change in the prism itself could result from modifications in the inlet or in the bay, and must be estimated from hydraulic considerations which have been described elsewhere (6).

Another observation is that, for large inlets, h_{cmax} is found to be nearly twice as much for inlets with jetties in comparison to those without jetties. At least a part of the reason for this is that, at many inlets where jetties exist, the channel is maintained at a required navigation depth by dredging. This effect is inherent in the coefficients of Eqs. 12 and 17 that are derived from measurements at inlets with jetties. Jetties also tend to confine the channel and thus the scouring capabilities of the flow is usually higher (4).

The depth of submarine pipeline burial must of course exceed h_{cmax}, since the latter is the depth to which a stable inlet on a sandy coast may be expected to scour. Unstable inlets are usually characterized by shoaling which means that scour depths will in general be less than h_{cmax}.

Conclusions

Design criteria for burial depths for submarine pipelines across tidal rivers, estuaries and inlets should be based on the rational considerations utilizing the tractive force theory approach rather than one based upon regime theory, whose applicability to cases involving tidal flow is tenuous at best, and is not recommended. For the case of shallow, alluvial tidal rivers and estuaries, a design case based upon a statistical analysis of real-time velocity records is suggested. Depth-averaged one- and two-dimensional hydrodynamic models, in which the boundary conditions are based on the real-time records, may be utilized to determine instantaneous depths and velocities at the location under consideration. The corresponding bed shear stresses can then be computed for peak velocities by assuming the existence of a logarithmic velocity distribution. In cases where cohesive bed material is encountered, a geotechnical analysis will, in general, be required to estimate the critical shear stress for erosion.

The tractive force approach is also applicable to tidal inlets, where, however, additional considerations must be given to channel stability under wave action in the mouth region. The gorge section of the channel is more stable then the mouth, and crossings at the mouth must be avoided. For design estimation of the maximum depth in a proposed inlet, a simple relationship based upon extensive field data is found to yield the maximum scour depth, given the spring tidal prism. For inlets that are to be modified, the estimated change in the spring tidal prism derived from hydraulic considerations can be used to predict the corresponding change in the maximum scour depth. It is found that small inlets, i.e., those with surface widths at the throat of less than 150 m, are far more susceptible to changes in the spring tidal prism, than larger inlets of widths greater than 150 m. It is noted that crossings usually occur at small inlets.

Acknowledgment

Helpful suggestions and comments by N. Ray are sincerely acknowledged. Preparation of this manuscript was made possible by support from the Florida Sea Grant College, NOAA, Department of Commerce.

Appendix I.—References

1. Arulanandan, K., Gillogley, E., and Tully, R., "Development of a Quantitative Method to Predict Critical Shear Stress and Rate of Erosion of Natural Undisturbed Cohesive Soils," *Technical Report GL-80-5,* U.S. Army Engineer Waterways Experiment Station, Vicksburg, Miss., July, 1980.
2. Blench, T., "Regime Theory Equations Applied to a Tidal River Estuary," *Proceedings of the International Association Hydraulics Research Convention,* Minneapolis, Minn., 1953, pp. 77–83.
3. Blench, T., *Mobile-Bed Fluviology,* The University of Alberta Press, Edmonton, Alberta, Canada, 1969.
4. Bruun, P., *Stability of Tidal Inlets: Theory and Engineering,* Elsevier Scientific Publishing Co., Amsterdam, The Netherlands, 1978.
5. Chantler, A. G., "The Applicability of Regime Theory to Tidal Watercourses," *Journal of Hydraulic Research,* Vol. 12, No. 2, 1974, pp. 181–192.
6. Dean, R. G., "Hydraulics of Inlets," *Report No. UFL/COEL-71/019,* Coastal and Oceanographic Engineering Laboratory, University of Florida, Gainesville, Fla, 1971.
7. Escoffier, F. F., "Hydraulics and Stability of Tidal Inlets," *GITI Report No. 13,* U.S. Army Corps of Engineers, Coastal Engineering Research Center, Fort Belvoir, Va, Aug., 1977.
8. "Fraser River Pipeline Crossing," Northwest Hydraulic Consultants Ltd., submitted to Gas Engineering Division, British Columbia Hydro and Power Company, Edmonton, Alberta, Canada, Feb., 1975.
9. Graham, D. S., "Hydrodynamic Criteria for Submarine Pipeline Burial Regime versus Rational Theory," *Pipelines in Adverse Environments,* ASCE, Vol. 1, 1979, pp. 103–123.
10. Graham, D. S., "A Design Method for Pipeline River Crossings," *Transportation Engineering Journal of ASCE,* Vol. 106, No. TE2, Proc. Paper 15261, Mar., 1980, pp. 141–153.
11. Graham, D. S., and Machemehl, J. L., "Appropriate Force Coefficients for Ocean Pipelines," presented at the February 3–7, 1980, American Society of Mechanical Engineers Energy Technology Conference and Exhibition, New Orleans, La. (Paper 80-Pet-61).
12. Graham, D. S., and Mehta, A. J., "Burial Design Criteria for Tidal River and Inlet Crossings," presented at the April 14–18, 1980, ASCE Convention and Exposition held at Portland, Oreg. (Preprint 80–033).
13. Harleman, D. R. F., "Tidal Dynamics in Estuaries, Part II: Real Estuaries," *Estuary and Coastline Hydrodynamics,* A. T. Ippen, ed., McGraw-Hill Book Co., Inc., New York, N.Y., 1966, pp. 522–545.
14. Harleman, D. R. F., and Lee, C. H., "The Computation of Tides and Currents in Estuaries and Canals," *Technical Bulletin No. 16,* Committee on Tidal Hydraulics, U.S. Army Engineer Waterways Experiment Station, Vicksburg, Miss., Sept., 1969.
15. Hayter, E. A., and Mehta, A. J., "Verification of Changes in Flow Regime due to Dike Breakthrough Closure," *Coastal Structures 79,* ASCE, 1979, pp. 729–749.
16. Henderson, F. M., *Open Channel Flow,* The MacMillan Co., New York, N.Y., 1966.
17. Jarrett, J. T., "Tidal Prism-Area Relationships," *GITI Report No. 3,* U.S. Army Corps of Engineers, Coastal Engineering Research Center, Fort Belvoir, Va., Feb., 1976.
18. Jonsson, I. G., "Wave Boundary Layers and Friction Factors," *Coastal Engineering,* ASCE, Vol. I, 1967, pp. 127–148.
19. Komar, P. D., and Miller, M. C., "Sediment Threshold under Oscillatory Waves," *Coastal Engineering,* ASCE, Vol. II, July, 1974, pp. 756–775.
20. Krishnamurthy, M., "Tidal Prism of Equilibrium Inlets," *Journal of the Waterway, Port, Coastal and Ocean Division,* ASCE, Vol. 103, No. WW4, Proc. Paper 13315, Nov., 1977, pp. 423–432.
21. Leopold, L. B., Wolman, M. G., and Miller, J. P., *Fluvial Processes in Geomorphology,* W. H. Freeman and Co., San Francisco, Calif., 1964.
22. Mason, C., "Regime Equations and Tidal Inlets," *Journal of the Waterways, Harbors and Coastal Engineering Division,* ASCE, Vol. 99, No. WW3, Proc. Paper 9952, Aug., 1973, pp. 393–397.

23. McDowell, D. M., and O'Connor, B. A., *Hydraulic Behaviour of Estuaries*, John Wiley and Sons, Inc., New York, N.Y., 1977.
24. Mehta, A. J., "Bed Friction Characteristics of Three Tidal Entrances," *Coastal Engineering*, No. 2, 1978, pp. 69–83.
25. Miller, M. C., McCave, I. N., and Komar, P. D., "Threshold of Sediment Motion under Unidirectional Currents," *Sedimentology*, Vol. 24, No. 4, Aug., 1977, pp. 507–527.
26. Myrick, R. M., and Leopold, L. B., "Hydraulic Geometry of a Small Tidal Estuary," *Professional Paper 422-B*, U.S. Geological Survey, Washington, D.C., 1963.
27. O'Brien, M. P., "Equilibrium Flow Areas of Inlets on Sandy Coasts," *Journal of the Waterways and Harbors Division*, ASCE, Vol. 95, No. WW1, Proc. Paper 6405, Feb., 1969, pp. 43–52.
28. Park, C. C., "World-Wide Variations in Hydraulic Geometry Exponents of Stream Channels: An Analysis of Some Observations," *Journal of Hydrology*, Vol. 33, No. 1/2, Mar., 1977, pp. 133–146.
29. Pillsbury, G. B., "Tidal Hydraulics," *Professional Paper 34*, U.S. Army Corps of Engineers, Washington, D.C., 1939, pp. 229–230.
30. Raudkivi, A. J., *Loose Boundary Hydraulics*, 2nd ed., Pergamon Press, Inc., Toronto, Canada, 1976.
31. Simons, D. B., and Richardson, E. V., "Flow in Alluvial Sand Channels," *River Mechanics*, H. W. Shen, ed., Vol. I, H. W. Shen Publisher, Fort Collins, Colo., 1971.
32. Wang, J. D., and Connor, J. J., "Mathematical Modeling of Near Coastal Circulation," *Report No. 200*, Ralph M. Parsons Laboratory for Water Resources and Hydrodynamics, Massachusetts Institute of Technology, Cambridge, Mass., Apr., 1975.
33. Ward, P. R. B., "The Transverse Distribution of Velocity in Estuary Flow," *Journal of Hydraulic Research*, Vol. 12, No. 2, 1974, pp. 253–274.
34. Winton, T. C., "Long and Short Term Stability of Small Tidal Inlets," thesis presented to the University of Florida, at Gainesville, Fla., in 1979, in partial fulfillment of the requirements for the degree of Master of Science.

Appendix II.—Notation

The following symbols are used in this paper:

A_c = inlet throat area;
A'_c = reference initial inlet throat area;
a = coefficient in Eq. 2;
a_1 = coefficient in Eq. 12;
a_2 = coefficient in Eq. 17;
a_3 = coefficient in Eq. 18;
a_4 = coefficient in Eq. 19;
a_i = inlet tidal amplitude;
b = exponent in Eq. 2;
c = coefficient in Eq. 3;
F_b = bed factor in Eq. 5;
f = exponent in Eq. 3;
h = mean river depth;
h_c = mean depth at inlet throat below mean water level;
h'_c = reference initial mean depth at throat;
h_{cmax} = maximum depth at throat;
h_0 = elevation above bed at which depth-mean velocity is equal to actual velocity;
h_r = regime depth;
h_s = scour depth in tidal river;

k = coefficient in Eq. 4;
k_s = Nikuradse's equivalent sand roughness of bed;
l = total length of inlet channel;
l_m = length of mouth section of inlet channel;
m = exponent in Eq. 4;
m_1 = exponent in Eq. 12;
m_2 = exponent in Eq. 17;
m_4 = exponent in Eq. 19;
n = total number of harmonic constituents of flow;
Q = discharge;
T = tidal period;
T_i = period of ith harmonic constituent of flow;
t = time;
$u(t)$ = time-dependent flow velocity in tidal river;
$u'(t)$ = aperiodic residual velocity in Eq. 1;
$\bar{u}$ = net flow-through velocity over tidal cycle;
u_m = depth-mean velocity;
u_{0i} = amplitude of ith harmonic constituent of flow;
u_* = shear velocity;
u_{*cr} = critical shear velocity for erosion;
w = surface river width;
w_c = surface width at inlet throat at mean water level;
z = z-factor in Eq. 6;
ρ = fluid (water) density;
Δh_{cmax} = change in h_{cmax} due to change $\Delta\Omega$ in spring tidal prism;
$\Delta\Omega$ = change in spring tidal prism due to inlet modification;
τ_{cr} = critical tractive stress for erosion;
τ_0 = bed shear stress or tractive stress;
ϕ_i = phase angle of ith harmonic constituent of the flow; and
Ω = spring tidal prism.

TRANSPORTATION
ENGINEERING JOURNAL

TECHNICAL NOTES

Note.—Discussion open until August 1, 1981. To extend the closing date one month, a written request must be filed with the Manager of Technical and Professional Publications, ASCE. This paper is part of the Transportation Engineering Journal of ASCE, Proceedings of the American Society of Civil Engineers, ©ASCE, Vol. 107, No. TE2, March, 1981.

Technical Notes

To provide a place within ASCE for publication of technical ideas that have not advanced, as yet, to the point where they warrant publication as a Proceedings paper in a *Journal,* the publication of Technical Notes was authorized by the Board of Direction on October 16–18, 1967, under the following guidelines:

1. An original manuscript and two copies are to be submitted to the Manager of Technical and Professional Publications, ASCE, 345 East 47th Street, New York, N.Y., 10017, along with a request by the author that it be considered as a Technical Note.
2. The two copies will be sent to an appropriate Technical Division or Council for review.
3. If the Division or Council approves the contribution for publication, it shall be returned to Society Headquarters with appropriate comments.
4. The technical publications staff will prepare the material for use in the earliest possible issue of the *Journal,* after proper coordination with the author.
5. Each Technical Note is not to exceed 4 pages in the *Journal.* As an approximation, each full manuscript page of text, tables, or figures is the equivalent of one-half a *Journal* page.
6. The Technical Notes will be grouped in a special section of each *Journal.*
7. Information retrieval abstracts and key words will be unnecessary for Technical Notes.
8. The final date on which a Discussion should reach the Society is given as a footnote with each Technical Note.
9. Technical Notes will not be included in *Transactions.*
10. Technical Notes will be included in ASCE's annual and cumulative subject and author indexes.

The manuscripts for Technical Notes must meet the following requirements:

1. Titles must have a length not exceeding 50 characters and spaces.
2. The author's full name, Society membership grade, and a footnote reference stating present employment must appear on the first page of the manuscript. Authors need not be Society members.
3. The manuscript is to be submitted as an original copy (with two duplicates) that is typed double-spaced on one side of 8-1/2-in. (220-mm) by 11-in. (280-mm) white bond paper.
4. All mathematics must be typewritten and special symbols must be properly identified. The letter symbols used must be defined where they first appear, in figures or text, and arranged alphabetically in an Appendix.—Notation.
5. Standard definitions and symbols must be used. Reference must be made to the lists published by the American National Standards Institute and to the *Authors' Guide to the Publications of ASCE.*
6. Tables must be typed double-spaced (an original ribbon copy and two duplicate copies) on one side of 8-1/2-in. (220-mm) by 11-in. (280-mm) paper. An explanation of each table must appear in the text.
7. Figures must be drawn in black ink on one side of 8-1/2-in. (220-mm) by 11-in. (280-mm) paper. Because figures will be reproduced with a width of between 3 in. (76 mm) to 4-1/2 in. (110 mm), the lettering must be large enough to be legible at this width. Photographs must be submitted as glossy prints. Explanations and descriptions must be made within the text for each figure.
8. References cited in text must be typed at the end of the Technical Note in alphabetical order in an Appendix.—References.
9. Dual units, i.e., U.S. Customary followed by SI (International System) units in parentheses, should be used throughout the paper.

Formulation of Parametric Measure of Track Quality

By A. E. Fazio,[1] M. ASCE and P. Olekszyk[2]

(Reviewed by the Urban Transportation Division)

Introduction

Track Quality can be generally defined as the ability of track to meet its functional requirements. Although a wide variety of parameters have been utilized to attempt to quantify the quality of track, none of these techniques have received industry-wide acceptance as being representative of the ability of track to properly support train movement (8).

A number of theoretical considerations have been published regarding the optimization problem with respect to Maintenance of Way resources, all of which presuppose both the existence of a parametric measure of track quality, and of a functional relationship which indicates how this parameter varies with applied maintenance and track usage (3,4).

This paper presents a summary analysis of the various attempts to formulate a "Track Quality Index" (TQI) for use by the railroad industry, and analyzes some specifics concerning the utilization of the loaded geometry of track in formulating a TQI.

Concept of a Track Quality Index

A general definition of a Track Quality Index (TQI) would be a parameter which provides a measure of the ability of railroad track to meet its functional requirements. More specifically, the TQI might be defined to measure certain probabilities such as of a derailment, or of lading damage, or could be constructed to be a measure of ride quality. Each of these criteria requires a consideration of the type of service to be provided. Thus track quality requires a consideration of the vehicle type and speed as well as an analysis of the track structure. The displacement and acceleration input received by the vehicle occurs at the wheel-rail interface and is determined by the loaded "surface" of the track.

[1]Project Engr., Conrail, Operations Planning, 6 Penn Center Plaza, Philadelphia, Pa. 19104.

[2]Chf., Analysis and Evaluation Div., Federal Railroad Administration, 400 Seventh Street, S.W., Washington, D.C. 20590.

Note.—Discussion open until August 1, 1981. To extend the closing date one month, a written request must be filed with the Manager of Technical and Professional Publications, ASCE. Manuscript was submitted for review for possible publication on June 4, 1980. This paper is part of the Transportation Engineering Journal of ASCE, Proceedings of the American Society of Civil Engineers, ©ASCE, Vol. 107, No. TE2, March, 1981. ISSN 0569-7891/81/0002-0245/$01.00.

Loaded surface (relative and absolute position of the running rails measured in both the horizontal and vertical planes) is a function of vehicle configuration and velocity as well as a wide variety of track components, e.g., ballast type and depth, tie spacing, etc. The TQI concept is an attempt to achieve a parametric summarization of the loaded track surface.

There are a wide variety of TQIs presently under investigation, with some receiving limited use in such applications as maintenance of way planning and quality control. These TQIs can be considered as belonging to one of three categories:

1. Physical characteristics of track structure, e.g., defective tie counts.—The rational for using this type of TQI is that it provides an indication of the behavior

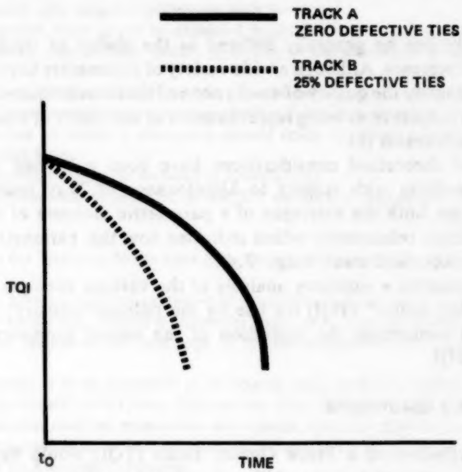


FIG. 1.—Change in Quality of Track with Time Immediately Subsequent to Surfacing

of track under service loads. The advantage of this type parameter lies primarily in its readily interpretated physical meaning. Unfortunately, the behavior of track structure is too complex and based in too many diverse factors to be accurately estimated by a TQI of this type.

2. Direct measurements of track strength, e.g., resistance to gage widening.—A direct measurement can be made of one or more aspects of track strength by recording the force required to achieve a given deflection. Performing such measurements with vehicles capable of automatically recording and processing data is a relatively new technique. A significant advantage of this type of TQI is that the imposed loadings can be made representative of that experienced by track during general freight service. Track strength alone, however, provides an incomplete estimation of vehicle behavior. Track gage 3 in. wide will result in a derailment regardless of the strength of the track.

3. Parameters derived from loaded geometry of track measured in both vertical and horizontal planes.—A variety of sophisticated track geometry cars are available to automatically measure and record track gage, alinement, cross level and warp (change in cross level). Each of these parameters has a certain frequency distribution over track. Investigations are underway regarding the suitability of utilizing statistics from these distributions as TQIs. An advantage of this approach is in the ease in which data may be collected, processed, and analyzed utilizing automatic track geometry cars. Two significant disadvantages of these

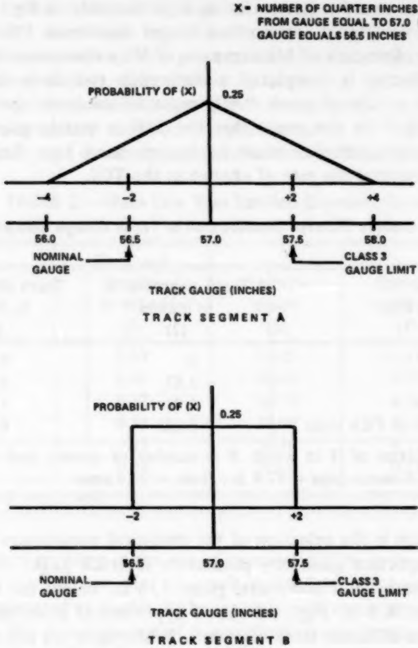


FIG. 2.—Hypothetical Probability Density Functions for Track Gage (1 in. = 25.4 mm)

type TQIs are that the loads at which the geometry is measured are often significantly lighter than the track receives in general freight service, and that many of the statistics utilized do not lend themselves to an easily understood physical interpretation.

Points of Consideration Regarding Geometric TQIs

Since geometric TQIs measure the loaded track surface directly, much intensive research has been directed towards their development. Professor Baluch, of

the Polish State Railways, has conducted a survey of much of this work (1). It is, however, questionable whether structural characteristics can be totally ignored in the formulation of a TQI. Consider two hypothetical tracks, both of which are identical in terms of their structure and the usage they receive except that track "A" contains no defective ties and in track "B" 25% of the ties are defective. If both tracks are surfaced simultaneously, a TQI which is strictly geometric will indicate the same (or nearly the same) quality for each track immediately subsequent to surfacing, i.e., it is possible that the two tracks will not differ appreciably in their loaded geometry if it is recorded by a track geometry car shortly after surfacing is performed. As Fig. 1 indicates, however, track "A" will "hold" its surface longer than track "B." If a TQI is to be of use in the allocation of Maintenance of Way resources, it is evident that at the time surfacing is completed a distinction (which is reflective of the longer maintenance life of track "A") must be made in the conditions of track "A" and "B." In the event that the TQI is purely geometric, the pertinent structural characteristics must be incorporated into the functional relationship which describes the rate of change in the TQI.

TABLE 1.—Values of Possible TQIs for Distribution of Track Gauge Shown in Figure 2

Possible TQI (1)	Track segment A, in inches (2)	Track segment B, in inches (3)
Mean	0	0
Variance	2.67	1.33
Ninety-ninth percentile	3.98	1.96
Probability of violation to FRA class 3	0.125	0

Note: Values are in terms of X in which X = number of quarter inch variation of gage from 57.0 in.; $X = 0$ means gage = 57.0 in.; (1 in. = 25.4 mm).

Another consideration is the selection of the statistical parameters to be used as TQIs, e.g., an important geometry parameter is track gage; the distance between the rail measured in a horizontal plane 5/8 in. below the top of rail. Standard track gage is 56.5 in. Figure 2 shows hypothetical probability density functions (pdfs) for two different track segments. While these are not necessarily actual pdfs for gage, they do provide an illustration of the care required in the development of geometric TQIs. Values of statistics which could be utilized as TQIs are shown for each distribution in Table 1.

Assume that both segments of track are class three according to the Federal Railroad Administration (FRA) safety standards. These standards permit up to 1 in. wide (57.5 in.) gage for class three track; any gage reading wider than this constitutes a safety violation. Both track segments have a mean gage of 57.0 in. In segment "B" there is zero probability of finding a value of track gage which violates the class three limit. In segment "A" there is a probability of 0.125 of a random gage observation exceeding in class limit.

Table 2 lists actual data recorded by the FRA's track geometry car on a main line test section. Because track geometry cars make discrete measurements the values shown are sample statistics, i.e., sample mean, sample standard

deviation, and the ninety-ninth percentile level in the sample. Segment length is included in order to provide an indication of sample size. In each segment data was recorded every foot. Of particular interest is the contrast in segments number 233 and 284. While both have essentially the same mean gage the measures of extremum (variance and ninety-ninth percentile) indicate that there are gage problems in segment 233.

The automatic track geometry cars are capable of measuring a number of different geometry parameters in addition to track gage including cross level and warp. Each of these geometry parameters has a density distribution over a segment of track which is defined by a number of statistical parameter, e.g., mean, standard deviation. In order to select an appropriate set of TQIs, it is necessary to determine which geometric and statistical parameters significantly affect vehicle behavior. One promising technique is to model the vehicle as a mathematical function which operates on an input signal, i.e., the track geometry (2,5). Other attempts at relating vehicle behavior to track geometry have utilized empirical data, e.g., derailments, in regression modeling (7).

TABLE 2.—Main Line Test Section Segment Sample Data

Segment number (1)	Segment length, in miles (2)	Sample mean (3)	Sample variance (4)	99% of sample (5)
187	0.69	57.02	0.17	57.47
199	0.61	56.98	0.27	57.55
233	0.17	56.99	0.31	57.85
268	0.31	57.00	0.16	57.39
284	0.15	56.99	0.17	57.36
306	0.30	57.00	0.21	57.38

Note: 1 mile = 1.61 km.

A significant problem in the use of geometric TQIs is the lack of a constitutive relationship for the behavior of track under load. Although track structure is often approximated as elastic, its behavior is both stress dependent and stresspath dependent. The development of a constitutive relationship for track is an area of intensive research (6,9).

Summary

It is apparent that any parametric measure of track quality should be able to be collected automatically and must:

1. Originate with a definition of track's performance requirements, e.g., minimize probability of derailment at speeds of up to fifty MPH.
2. Incorporate vehicle dynamics in ascertaining track's ability to satisfy its performance requirements.
3. Provide a measure of the probability of a discrete failure in track performance in addition to an overall measure of track quality. This will require two families

of TQIs, one of which is a measure of extremum and the other which measures central tendency of a distribution.

4. Somewhere, either in the TQI itself or in the functional relationship describing its rate of change, incorporate pertinent physical characteristics of the track which are not capable of being recorded automatically.

Efforts to develop such a parameter in Europe as well as in this country are underway. Conrail and the FRA currently have a cooperative program underway in an attempt to formulate such a number based on automatic track geometry car data.

Appendix.—References

1. Balunch, et al., "Methodology for Quantification and Means of Improvement of Track Structure Condition with a Particular Interest to its Heterogeneity," Polish People's Republic Institute for Railway Research, Warsaw, Poland, 1977.
2. Corbin, J. C., and Kaufman, W. M., "Classifying Track by Power Spectral Density," *Mechanics of Transportation Suspension Systems, Proceedings of the Annual Meeting of the American Society of Mechanical Engineers*, Dec., 1975.
3. Fazio, A. E., and Prybella, R., "Development of an Analytical Approach to Track Maintenance Planning," *Transportation Research Record 744*, Railroad Track and Facilities Transportation Research Board.
4. Friesz, T. L., and Fernandez-Larranaga, "A Model of Optimal Transport Maintenance with Demand Responsiveness," *Transportation Research Board Publication*.
5. Rinehart, R. E., "Locomotive Response to Random Track Surface Irregularities," Proceedings of the Annual Meeting of the American Society of Mechanical Engineers, 78-WA/RT-12, Dec., 1978.
6. Selig, E. T., et al., "A theory for Track Maintenance Life Prediction," *United States Department of Transportation Report DPB-50/79/22*, Aug., 1979.
7. Simpson, W. W., "Southern Railway Research," *American Railway Engineering Association Bulletin No. 673*, June, 1979, pp. 372–391.
8. Tayabji, S. D., and Thompson, M., "Considerations in the Analysis of Conventional Railway Track Support System," *Transportation Engineering Journal of ASCE*, ASCE, Vol. 103, No. TE2, Proc. Paper 12792, Mar., 1977, pp. 280–291.
9. Zarembski, A. M., and Choros, J., *Laboratory Investigation of Track Gauge Widening*, Report No. R-395, Association of American Railroads, Sept., 1979.

TRANSPORTATION ENGINEERING JOURNAL

DISCUSSION

Note.—This paper is part of the Transportation Engineering Journal of ASCE, Proceedings of the American Society of Civil Engineers, ©ASCE, Vol. 107, No. TE2, March, 1981. ISSN 0569-7891/81/0002-0253/$01.00.

DISCUSSIONS

Discussions may be submitted on any Proceedings paper or technical note published in any *Journal* or on any paper presented at any Specialty Conference or other meeting, the *Proceedings* of which have been published by ASCE. Discussion of a paper/technical note is open to anyone who has significant comments or questions regarding the content of the paper/technical note. Discussions are accepted for a period of 4 months following the date of publication of a paper/technical note and they should be sent to the Manager of Technical and Professional Publications, ASCE, 345 East 47th Street, New York, N.Y. 10017. The discussion period may be extended by a written request from a discusser.

The original and three copies of the Discussion should be submitted on 8-1/2-in. (220-mm) by 11-in. (280-mm) white bond paper, typed double-spaced with wide margins. The length of a Discussion is restricted to two *Journal* pages (about four typewritten double-spaced pages of manuscript including figures and tables); the editors will delete matter extraneous to the subject under discussion. If a Discussion is over two pages long it will be returned for shortening. All Discussions will be reviewed by the editors and the Division's or Council's Publications Committees. In some cases, Discussions will be returned to discussers for rewriting, or they may be encouraged to submit a paper or technical note rather than a Discussion.

Standards for Discussions are the same as those for Proceedings Papers. A Discussion is subject to rejection if it contains matter readily found elsewhere, advocates special interests, is carelessly prepared, controverts established fact, is purely speculative, introduces personalities, or is foreign to the purposes of the Society. All Discussions should be written in the third person, and the discusser should use the term "the writer" when referring to himself. The author of the original paper/technical note is referred to as "the author."

Discussions have a specific format. The title of the original paper/technical note appears at the top of the first page with a superscript that corresponds to a footnote indicating the month, year, author(s), and number of the original paper/technical note. The discusser's full name should be indicated below the title (see Discussions herein as an example) together with his ASCE membership grade (if applicable).

The discusser's title, company affiliation, and business address should appear on the first page of the manuscript, along with the *Proceedings* paper number of the original paper/technical note, the date and name of the *Journal* in which it appeared, and the original author's name.

Note that the discusser's identification footnote should follow consecutively from the original paper/technical note. If the paper/technical note under discussion contained footnote numbers 1 and 2, the first Discussion would begin with footnote 3, and subsequent Discussions would continue in sequence.

Figures supplied by the discusser should be designated by letters, starting with A. This also applies separately to tables and references. In referring to a figure, table, or reference that appeared in the original paper/technical note use the same number used in the original.

It is suggested that potential discussers request a copy of the *ASCE Authors' Guide to the Publications of ASCE* for more detailed information on preparation and submission of manuscripts.

ENVIRONMENTAL CONSIDERATIONS IN HIGHWAY PLANNING[a]

Discussion by William R. Green[3]

The authors are to be commended for their objective and constructive analysis of a subjective and sometimes emotional subject. The writer would like to comment based on his experience as Caltrans' Chief of Office of Planning and Design, which had a primary responsibility for developing statewide policies and instructions to cope with the rapid changes in planning procedures brought about by NEPA, CEQA (California Environmental Quality Act), and the numerous implementing regulations during the decade of the 1970s. It is hoped that these comments will supplement the authors' article by adding insight from a different perspective.

First, it can't be overemphasized that the regulations promulgated as a result of the passage of NEPA and CEQA had a significant impact on the planning and design of public works projects *and* private development.

The impact was not caused, in Caltrans' case at least, by the imposition of totally new requirements which we have never done before. Caltrans had been a leader in environmental and aesthetic considerations long before NEPA, and had received national awards and recognition such as the Annual Parade Magazine award for Route 88 and many FHWA prizes. A unit in Caltrans called Community and Environmental Factors Unit (CEFU), which preceded the present Office of Environmental Planning, made significant contributions to this effort in the 1960s.

Rather, it was because the new requirements were applied retroactively to projects which, as the authors noted, had been in the planning and design pipeline for many years, with major decisions such as route location, width of median and R/W, interchange location, etc., already made. In many cases, most of the R/W had already been purchased, and as the authors noted, development constructed up to the future R/W lines, effectively foreclosing any other location.

It shouldn't be surprising then, that the detailed study, consideration, and documentation of location or scope alternatives required for these projects were seen as completely unrealistic (such as "no project" for a 5-mile gap in 50 miles of existing full Interstate freeway) and met with something less than overwhelming enthusiasm by Caltrans engineers. The large amount of work required, the time delay, and considerable expense, both in staff costs and inflated construction costs caused by delay, were seen largely as paperwork exercises with little or no purpose other than documentation to be served; particularly because of what was felt to be a pretty good track record for environmental considerations in making the past major decisions.

---

[a]July, 1980, by John J. Meersman and Leonard Ortolano (Proc. Paper 15525).

[3]Chf. Office of Office Engr., California Dept. of Transportation, 1120 N St., Sacramento, Calif.

As the authors noted, one of the most important factors in coping with any change is people's attitude. If the reasons for the change are understood, and they're seen as necessary (which is management's job to accomplish), a positive attitude can result which will minimize problems. In Caltrans, as we have progressed beyond pipeline projects and been able to apply environmental requirements to new projects, where changes can be made to reduce or eliminate impacts in a *timely* fashion, the desirability and advantage of full consideration of all social-environmental-economic (S-E-E) effects is much more apparent, and attitudes are correspondingly more positive.

The desirability of broad-based public participation from the early planning stages is noted. The writer fully agrees and feels that the one formal hearing now required at the time of circulation of the DEIS should be a mere formality if planners have done their job. There should be no surprises at this time if the concerns and positions of affected agencies and groups have been identified during the planning process. Unfortunately, it is an almost universal human trait to hope that potential or identified problems will disappear if ignored long enough, and Caltrans is not unique in having its share of human beings. It was suggested that Project Development teams would benefit from a broad-based membership. Caltrans team memberships can be quite broad, depending on the project, and have included Sierra Club representatives, Highway Patrol, Citizen Aesthetic Advisors, Department of Fish and Game, FHWA, and so on. It has been our experience that valuable contributions can be made by these outside members.

The authors stated that the Caltrans Action Plan was adopted in July, 1973 and, therefore, their survey was based on events after November, 1973 when the Action Plan was in full effect. In the writer's opinion, it took much longer than that for it to become well understood, since a number of new requirements were included in the plan.

Landscaping seems to be considered by the authors as environmental mitigation. Caltrans regards landscaping, which it has provided for many years, more as the last element of a complete project.

The authors have provided considerable insight into this subject and are to be commended on an interesting approach.

[illegible]

[illegible]

1. [illegible]

2. [illegible]

3. [illegible]

4. [illegible]

5. [illegible]

6. [illegible]

7. [illegible]

8. [illegible]

9. [illegible]

10. [illegible]

11. [illegible]

12. [illegible]

13. [illegible]

14. [illegible]

Technical Papers

Original papers should be submitted in triplicate to the Manager of Technical and Professional Publications, ASCE, 345 East 47th Street, New York, N.Y. 10017. Authors must indicate the Technical Division or Council, Technical Committee, Subcommittee, and Task Committee (if any) to which the paper should be referred. Those who are planning to submit material will expedite the review and publication procedures by complying with the following basic requirements:

1. Titles must have a length not exceeding 50 characters and spaces.

2. The manuscript (an original ribbon copy and two duplicate copies) should be double-spaced on one side of 8-1/2-in. (220-mm) by 11-in. (280-mm) paper. Three copies of all figures and tables must be included.

3. Generally, the maximum length of a paper is 10,000 word-equivalents. As an *approximation*, each full manuscript page of text, tables or figures is the equivalent of 300 words. If a particular subject cannot be adequately presented within the 10,000-word limit, the paper should be accompanied by a rationale for the overlength. This will permit rapid review and approval by the Division or Council Publications and Executive Committees and the Society's Committee on Publications. Valuable contributions to the Society's publications are not intended to be discouraged by this procedure.

4. The author's full name, Society membership grade, and a footnote stating present employment must appear on the first page of the paper. Authors need not be Society members.

5. All mathematics must be typewritten and special symbols must be identified properly. The letter symbols used should be defined where they first appear, in figures, tables, or text, and arranged alphabetically in an appendix at the end of the paper titled Appendix.—Notation.

6. Standard definitions and symbols should be used. Reference should be made to the lists published by the American National Standards Institute and to the *Authors' Guide to the Publications of ASCE*.

7. Figures should be drawn in black ink, at a size that, with a 50% reduction, would have a published width in the *Journals* of from 3 in. (76 mm) to 4-1/2 in. (110 mm). The lettering must be legible at the reduced size. Photographs should be submitted as glossy prints. Explanations and descriptions must be placed in text rather than within the figure.

8. Tables should be typed (an original ribbon copy and two duplicates) on one side of 8-1/2-in. (220-mm) by 11-in. (280-mm) paper. An explanation of each table must appear in the text.

9. References cited in text should be arranged in alphabetical order in an appendix at the end of the paper, or preceding the Appendix.—Notation, as an Appendix.—References.

10. A list of key words and an information retrieval abstract of 175 words should be provided with each paper.

11. A summary of approximately 40 words must accompany the paper.

12. A set of conclusions must end the paper.

13. Dual units, i.e., U.S. Customary followed by SI (International System) units in parentheses, should be used throughout the paper.

14. A practical applications section should be included also, if appropriate.